UNIVERSITÀ DEGLI STUDI DI TRENTO

DOTTORATO DI RICERCA IN MATEMATICA

XVI CICLO

Silvia Rensi

Analyzing complex choices through generalized spectral analysis: preference formation.

Relatore
Prof. Andrea Caranti

Preface

This Thesis deals principally with an application of the so–called generalized spectral analysis to a problem of economic science: preference formation. The needed mathematical framework is based on group representation theory, isotypic projections, spectral analysis on group.
The basic purpose of the whole is essentially efficient data analysis; this is obtained through the application of generalized spectral analysis (a generalization of classical spectral analysis, the one used in time series analysis, for example) to group theory . This kind of approach is very strong and allows to detect in data particular effects which are not appreciable from a direct analysis of data.
Our work is modeled on some recent results of M. E. Orrison and B. L. Lawson who applied generalized spectral analysis to political science; their work provides a sistematic way to detect influential coalitions in a political voting process.

In Chapter 1 we recall some mathematical background in order to reach a deeper comprehension of spectral analysis and its applications. The content of the chapter develops through some basic notion on group representation theory, a brief account on the representation theory of the symmetric group, the concept of Fourier transform and the techniques of the Fast Fourier transform and finally spectral analysis and its applications.

Section 1.1 is given up to a brief account on group representation theory; it is nothing else that a recall of the concept of representations, characters, irreducible representations and so on.

Section 1.2 deals with isotypic projections. Spectral analysis applied to groups is strongly based on the fact that any representation of a finite group can be decomposed in what is the so–called *isotypic decomposition*.

In section 1.3 we recall briefly the construction of all irreducible representations of the symmetric group and the decomposition of the permutation module.

Section 1.4 is given up to Fourier transform on groups. Originally discovered by Gauss and later made famous by Cooley and Tukey, the Fast Fourier transform may be viewed as an algorithm which efficiently computes the Fourier Transform. Recently, there has developed a growing literature related to the construction of algorithms which generalize the FFT from the point of view of the theory of group representations. These sort of generalizations are "natural" as mathematical constructs, but in point of fact they too have been motivated by applications, such as, as in our context, efficient data analysis.

Section 1.5 is the core of Chapter 1 and provides an account on generalized spectral analysis. Spectral analysis is a non–model based approach to data analysis, formulated in general group theoretic setting by Diaconis; it extends the classical spectral analysis of time series. The idea of spectral analysis is that often data has natural symmetries, encapsulated in the existence of a symmetry group for the domain of the data. The organizing principle of spectral analysis is the understanding of data trough its decomposition according to these symmetries.
If X is a finite set, G a group acting on X and $L(X)$ the vector space of complex–valued functions on G, then $L(X)$ may be decomposed as an orthogonal direct sum of G-invariant subspaces

$$L(X) = V_1 \oplus \cdots \oplus V_h,$$

called the isotypic decomposition. Spectral analysis takes the form of computing the projections of the data onto these subspaces and judging which projections are significant. The use of generalized FFTs for spectral analysis is the efficient computation of the projections.
Of particular interest is the analysis of ranking problems. Data sometimes come in the form of rank of preferences. Most anyone who analyzes such data looks at simple averages, such as the proportion of times each item was ranked first and the average rank for each item. There are **first order**

statistics: they are linear combinations of the number of times item i was ranked in position j. There are also natural **second order statistics**, based on the number of times items i and j are ranked in position k and l. Similarly there are third and **higher order statistics** of various type.
Diaconis underlines the crucial point od spectral analysis: "*a basic idea of data analysis is this: if you've found some structure, take it out and look at what is left. Thus, to look at second order statistic it is natural to substract away the observed first order structure. This leads to a natural decomposition of the original data into orthogonal pieces*".

Chapter 2 is given up to the exposition of some recent works of M. E. Orrison and B. L. Lawson (see [22], [23] and [24]) on noncommutative harmonic analysis applied to political voting.

Diaconis extended the classical spectral analysis of time series to a non–time series subject, for the analysis of discrete data which has a noncommutative structure. New efforts have been made in order to apply spectral analysis to a non–time series subject in political science, above all in the analysis of voting. Spectral analysis has been already used in political science to identify cycles in time series data.

Orrison and Lawson introduced a generalization of spectral analysis as a new instrument for political scientist; they used the powerful machinery of spectral analysis to analyze political voting data. In particular, they analyzed votes of the nine judges of the United States Supreme Court and detected influential coalitions. With this theory political scientist can use spectral analysis as a method for identifying substantively important dynamics in politics, rather then just as a diagnostic tool.
The idea followed by Orrison and Lawson is to consider political voting data as elements of a mathematical framework; then the features of that framework can be used to work out natural interpretations of the data. The mathematical framework corresponding to voting data has many components, each of which encapsulates information on particular *coalition effects*; the decomposition of data with respect to these components provides the identification of **influential coalitions**.

Chapter 3 is entirely given up to our contribution to the Thesis, with an application of generalized spectral analysis to preference formation.

Our context can be summarized as follows. We interpret the decision to vote for a party as a process of delegation to decision makers having a simplified system of preferences. Each person in a population votes for the political party that place priority on one or more issues that they consider important. On the basis of a survey on preferences of population, we have simulated a delegation procedure which chart the selection process of a particular party. Making use of noncommutative harmonic analysis, we decomposed the delegation function and isolated the effect of a particular affinity, or a combination of either the pair of items that characterize a party.

To be more precise, our construction bases itself on these considerations. Individuals facing a choice are often not able to make a full comparison between alternatives. Even if they are able to pin down their preferences for certain characteristics of an object (for instance, a car), they would probably be able to compare only a few of them. In the case of a car, one person would take into account room and safety, while somebody else's order ranking would be based on speed and acceleration. We can interpret this evaluation imagining that our "complete" selves delegate choices to a sort of simplified self.

In public choice theory, political parties present themselves as decision makers committed to following a given preference order when faced with future choices. Parties collect delegations from people having similar preferences. Traditionally parties have a complete system of preferences and they collect a delegation from the people having an order of preferences "not far" from the one expressed by the party.
Here instead of following this traditional path, we adopt a similar approach to the one presented in "car choice". **We describe parties as simplified systems of preferences and the process of delegation as giving the power of choice to parties that correspond to this simplified preference order**.

Given that parties compete to attract electors in a simplified preference space, the distribution of preferences will depend on the way preferences are simplified. If, for instance, parties simplify things proposing a couple of items to which they attach more importance, it could be that the items chosen complement themselves well, being able to attract a large share of voters, or alternatively the two items could reciprocally depress their power of attraction. When facing a simplified set of options, the right combination could be of fundamental importance.

Acknowledgements

I would like to thank my advisor, prof. Andrea Caranti, who introduced me to the variegated world of the subject of this Thesis.

I would like to give my heartfelt thanks to prof. Enrico Zaninotto, who helped me for almost all the period of my work and has been supportive and helpful through various difficult times.

Finally, I would like to thank with never ending love my family.

Contents

1
Spectral analysis and group representation theory

As explained in the Preface and examined carefully in Chapter 3, our work want to be an application of the so–called generalized spectral analysis in the field of economic science.

Spectral analysis is a non–model based approach to data analysis, formulated in general group theoretic setting by Diaconis (see [9], [10]); it extends the classical spectral analysis of time series. The idea of spectral analysis is that often data has natural symmetries, encapsulated in the existence of a symmetry group for the domain of the data. The organizing principle of spectral analysis is the understanding of data trough its decomposition according to these symmetries.

In order to reach a deeper comprehension of spectral analysis and its applications, we need some theoretical backgrounds. In this Chapter first of all we will recall some basic notation and terminology of group representation theory. It will follow a brief account on the representation theory of the symmetric group: we need the construction of all irreducible representations of

S_n and the decomposition of the permutation module. In our exposition it will be of great importance the concept of Fourier transform and the techniques of the Fast Fourier Transform. The core of the chapter will be spectral analysis and its applications.

1.1 Representation theory

We recall some general terminology and notation about group representation theory that will be useful in the following. Representation theory can be couched in terms of matrices or in the language of modules; we will occasionally consider both approaches. We will follow Serre (see [42]).

Let G be a finite group and V a finite dimensional vector space over $\mathbb{C}$. Let $GL(V)$ be the group of automorphisms of V. A **representation** of G is a group homomorphism $\rho : G \longrightarrow GL(V)$. If the homomorphism ρ is understood, then we also say that V is a representation of G. In the language of modules, V is also called a G–**module**. We call $d_\rho := \dim V$ the **degree** of the representation ρ.

Two representations ρ_1 and ρ_2 of a group G on V_1 and V_2 are said to be **isomorphic** if there exists an isomorphism $T : V_1 \longrightarrow V_2$ such that $\rho_1(s) = T^{-1} \circ \rho_2(s) \circ T$, for all $s \in G$. In other words, in matrices language, two representations are isomorphic if they differ only by a change of basis, i.e. there exists an invertible matrix A such that $\rho_1(s) = A^{-1}\rho_2(s)A$.

The **character** of ρ is the function $\chi : G \longrightarrow \mathbb{C}$ where $\chi(s)$ is the usual trace of $\rho(s)$. Note that the character of a representation of G is constant on the conjugacy classes of G.

A subspace W of V is **invariant** if $\rho(s)(w) \in W$, for all $s \in G, w \in W$. A representation is said to be **irreducible** if it contains no non–trivial invariant subspaces. If $C_1, \ldots, C_h$ are the distinct conjugacy classes of G, then there are h distinct irreducible representations $W_1, \ldots, W_h$ of G (up to isomorphisms). Irreducible representations are the fundamental blocks of all representations of a finite group. More precisely, any representation is isomorphic to a direct sum of irreducible representations.

1.2 Isotypic projections

As we will see, spectral analysis applied to group representation theory is strongly based on the fact that any representation of a finite group can be decomposed in what is the so–called *isotypic decomposition*. Serre (see [42]) gives a complete description of the decomposition, while Diaconis and Rockmore (see [12]) and Maslen, Orrison and Rockmore (see [29]) provide efficient algorithms for the computation of isotypic projections.

Let G be a finite group and $X = \{x_1, \dots x_n\}$ a finite set. Suppose that G acts transitively on X; X is called a *homogeneous space* for G. Let $L(X)$ denote the vector space of all complex-valued functions on X. Then $L(X)$ naturally admits a representation ρ of G defined by

$$\begin{aligned} &\rho : G \longrightarrow GL(L(X)) \\ &\rho(s) = \rho_s \end{aligned}$$

where

$$\begin{aligned} &\rho_s : L(X) \longrightarrow L(X) \\ &\rho_s(f)(x) = f(s^{-1}x) \end{aligned}$$

for each $s \in G$, $x \in X$ and $f \in L(X)$.

The vector space $L(X)$ has a natural basis $\{\delta_x\}_{x \in X}$, where

$$\delta_x(x') = \begin{cases} 1 & \text{if } \ x = x' \\ 0 & \text{otherwise.} \end{cases}$$

We will refer to $\{\delta_x\}_{x \in X}$ as the *delta basis* of $L(X)$. Note that $\dim L(X) = |X| := d_X$. By choosing a basis for $L(X)$, we may identify each linear transformation on $L(X)$ with a $d_X \times d_X$ matrix. Thus, we will assume that each linear transformation on $L(X)$ is written as a matrix with respect to the delta basis of $L(X)$. In particular, if $s \in G$, then ρ_s corresponds to a $d_X \times d_X$ matrix with one 1 in each row and column and zeros elsewhere.

The representation ρ obtained by $L(X)$ is a permutation representation of G. We recall the following

DEFINITION 1.2.1 *Let G be a finite group acting on a finite set X and V a vector space with a basis $\{e_x\}_{x\in X}$ indicized by the elements of X. The* ***permutation representation*** *associated to X is the representation φ defined by $\varphi : G \longrightarrow GL(V)$, $\varphi(s) = \varphi_s$, where*

$$\varphi_s : V \longrightarrow V$$
$$\varphi_s(e_x) = e_{sx}$$

for each $s \in G$ and $x \in X$.

Now,

$$\begin{aligned} \rho_s(\delta_x) : X &\longrightarrow \mathbb{C} \\ x' &\mapsto \delta_x(s^{-1}x') \end{aligned}$$

but

$$\delta_x(s^{-1}x') = \begin{cases} 1 & \text{if } \ x = s^{-1}x' \\ 0 & \text{otherwise} \end{cases} = \begin{cases} 1 & \text{if } \ sx = x' \\ 0 & \text{otherwise} \end{cases} = \rho_{sx}(x'),$$

which shows indeed that ρ is a permutation representation of G.

We recall that, if $X = \{x_1, \ldots, x_n\}$ is a finite set, for any $x_i \in X$, $\mathrm{Stab}(x_i)$ denote the stabilizer of x_i in G, that is the subgroup of elements of G which fix x_i. The representation ρ of G naturally defined by $L(X)$ is precisely the permutation representation of G on the quotient space $G/\,\mathrm{Stab}(x_i)$, for any i. Equivalently, it is the representation obtained by inducing the trivial representation from $\mathrm{Stab}(x_i)$ to G.

As a representation space for G, $L(X)$ has a basis indipendent decomposition into G-invariant subspaces, known as **isotypic subspaces**. Following Serre (see [42]), the so–called **isotypic decomposition** of $L(X)$ may be explained as follows.

Let $\rho_1, \ldots, \rho_h$ be a complete set of non-isomorphic irreducible representations of G and let $\chi_1, \ldots, \chi_h$ be the corresponding characters. $L(X)$ is a representation space for G, so can be decomposed into a direct sum of irreducible G-invariant subspaces. That is

$$L(X) = \bigoplus_{j=1}^{m} U_j.$$

where $\rho(s)U_j = U_j$, for each $U_j \neq 0$, and no trivial subspace of U_j has this property. For each $j = 1, \ldots, m$, $\rho(s)$ restricted to U_j gives an irreducible representation of G. In this decomposition of $L(X)$ there will be isomorphic copies of U_j, so we can define V_i as the subspace of $L(X)$ given by the direct sum over all U_j which define representations isomorphic to ρ_i, for each $i = 1, \ldots, h$. So the **isotypic decomposition** of $L(X)$ is

$$L(X) = \bigoplus_{i=1}^{h} V_i, \tag{1.1}$$

where V_i is called the i-**isotypic subspace** of ρ. (Note that some of the V_i may be zero).

Given an arbitrary $f \in L(X)$, we may compute the projection of f onto each isotypic subspace of $L(X)$. So, if $f \in L(X)$, f may be written uniquely as

$$f = f_1 + \cdots + f_h, \tag{1.2}$$

where f_i is called the **isotypic projection** of f onto the isotypic subspace V_i.

There is a classical theorem which allows to calculate the projection of a group representation V onto its isotypic subspaces (Serre [42], theorem no. 8, pag. 21).

THEOREM 1.2.1 (see Serre, [42]) *Let G be a finite group and $\varphi : G \longrightarrow GL(V)$ a representation of G. Let $\varphi_1, \ldots, \varphi_h$ be a complete set of non-isomorphic irreducible representations of G and $\sigma_1, \ldots, \sigma_h$ the corresponding characters. Let $V = V_1 \oplus \cdots \oplus V_h$ be the isotypic decomposition of V. For*

each $i = 1, \ldots, h$, define p_i as the projection of V on V_i.
Then

$$p_i = \frac{\deg(\varphi_i)}{|G|} \sum_{s \in G} \overline{\sigma_i(s)} \varphi(s) \tag{1.3}$$

for each $i = 1, \ldots, h$, where $\overline{\sigma_i(s)}$ is the conjugate of $\sigma_i(s)$.

In our particular case, where the representation is the permutation representation ρ obtained by $L(X)$, more can be said.

THEOREM 1.2.2 *Let G be a finite group acting on a finite set X. Let ρ be the associated permutation representation of G in $L(X)$, let $\rho_1, \ldots, \rho_h$ be a complete set of non-isomorphic irreducible representations of G and $\chi_1, \ldots, \chi_h$ the corresponding characters.*
For each $f \in L(X)$, let $f_1 + \cdots + f_h$ be the isotypic decomposition of f.
Then

$$f_i(x) = \frac{\deg(\rho_i)}{|G|} \sum_{s \in G} \chi_i(s) f(sx) \tag{1.4}$$

for each $x \in X$ and $i = 1, \ldots, h$.

Proof. We calculate

$$\begin{aligned}
f_i(x) = p_i(f(x)) &= \\
&= \frac{\deg(\rho_i)}{|G|} \sum_{s \in G} \overline{\chi_i(s)} \left(\rho(s) f\right)(x) = \\
&= \frac{\deg(\rho_i)}{|G|} \sum_{s \in G} \overline{\chi_i(s)} f(s^{-1}x) = \\
&= \frac{\deg(\rho_i)}{|G|} \sum_{s \in G} \chi_i(s^{-1}) f(s^{-1}x) = \\
&= \frac{\deg(\rho_i)}{|G|} \sum_{g \in G} \chi_i(g) f(gx).
\end{aligned}$$

where we used the elementary property $\overline{\chi_i(s)} = \chi_i(s^{-1})$.

□

1.3 Representations of the symmetric group

In this section we recall briefly the construction of all irreducible representations of the symmetric group S_n. Sagan (see [41]) gives a very nice exposition of the construction and James (see [18]) and James and Kerber (see [19]) provides an encyclopedic account with many references.

We know that the number of irreducible representations of S_n is equal to the number of conjugacy classes of S_n (see Serre [42]), which is also the number of partions of n. It is not obvious how to associate an irreducible representation of S_n with each partition $\lambda = (\lambda_1, \dots, \lambda_k)$ of n, but is quite easy to find a corresponding group S_λ that is an isomorphic copy of $S_{\lambda_1} \times S_{\lambda_2} \times \cdots S_{\lambda_k}$ inside S_n. The right number of irreducible representations of S_n may be produced by inducing the trivial representation on each S_λ up to S_n.

1.3.1 Tableaux, tabloids and Young subgroups

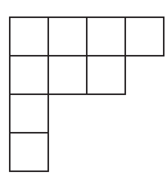
Let $\lambda \vdash n$ be a partition of n; this means that

$$\lambda = (\lambda_1, \dots, \lambda_k)$$

with $\lambda_1 \geq \cdots \geq \lambda_k > 0$ and $\lambda_1 + \cdots + \lambda_k = n$. In this case $k = h(\lambda)$ is called the **length** of λ. We can visualize λ as follows.

DEFINITION 1.3.1 *Let $\lambda = (\lambda_1, \dots, \lambda_k) \vdash n$ be a partition of n. A* ***Young diagram of shape*** *λ is a left-justified array of square boxes with λ_i boxes in row i, with $1 \leq i \leq k$.*

For example, the Young diagram of shape $(4, 3, 1, 1)$ is



DEFINITION 1.3.2 *A **Young tableau** t **of shape** λ is a Young diagram of shape λ with integers $1, \ldots, n$ placed without repetition in its boxes.*

For example, a Young tableau of shape $(4, 3, 1, 1)$ is

$$t = \begin{array}{|c|c|c|c|} \hline 3 & 1 & 2 & 4 \\ \hline 5 & 6 & 8 \\ \cline{1-3} 7 \\ \cline{1-1} 9 \\ \cline{1-1} \end{array}$$

Obviously there are $n!$ tableaux of a fixed shape.

DEFINITION 1.3.3 *Two Young tableaux t_1 and t_2 of shape λ are said to be **equivalent** if they differ only by permuting the entries within a given row. An equivalence class of Young tableaux of a fixed shape is called a **tabloid** of the same shape.*

For example, two equivalent Young tableaux of shape $(4, 3, 1, 1)$ are

$$t_1 = \begin{array}{|c|c|c|c|} \hline 1 & 2 & 3 & 4 \\ \hline 5 & 6 & 7 \\ \cline{1-3} 8 \\ \cline{1-1} 9 \\ \cline{1-1} \end{array}, \qquad t_2 = \begin{array}{|c|c|c|c|} \hline 2 & 1 & 3 & 4 \\ \hline 7 & 5 & 6 \\ \cline{1-3} 8 \\ \cline{1-1} 9 \\ \cline{1-1} \end{array}.$$

We may think of a tabloid as a tableau with unordered row entries. A given tabloid is denoted by forming the representative Young tableau and removing the internal vertical lines. For example, the tabloid of shape $(4, 3, 1, 1)$

$$\{t\} = \begin{array}{|cccc|} \hline 1 & 2 & 3 & 4 \\ \hline 5 & 6 & 7 \\ \cline{1-3} 8 \\ \cline{1-1} 9 \\ \cline{1-1} \end{array}$$

represents the equivalence class of the t_1 and t_2 of the previous example.
If $\lambda = (\lambda_1, \ldots, \lambda_k) \vdash n$, then the number of tableaux in any given equivalence class is $\lambda_1!\lambda_2! \cdots \lambda_k! = \lambda!$. Thus the number of tabloids of shape λ is just $\frac{n!}{\lambda!}$.

Now we wish to associate with a partition λ a subgroup of S_n.

DEFINITION 1.3.4 *Let $\lambda = (\lambda_1, \ldots, \lambda_k) \vdash n$ be a partition of n. The* ***Young subgroup of*** S_n *corresponding to λ is*

$$S_\lambda = S_{\{1,2,\ldots,\lambda_1\}} \times S_{\{\lambda_1+1,\lambda_1+2,\ldots,\lambda_1+\lambda_2\}} \times \cdots \times S_{\{n-\lambda_k+1,n-\lambda_k+2,\ldots,n\}}$$

where $S_{\{i,j,\ldots,l\}}$ means the subgroup of S_n permuting only the integers in the brackets.

These subgroups are named in honor of the Reverend Alfred Young, who was among the first to construct the irreducible representations of S_n. For example,

$$\begin{aligned} S_{(3,3,2,1)} &= S_{\{1,2,3\}} \times S_{\{4,5,6\}} \times S_{\{7,8\}} \times S_{\{9\}} \\ &\cong S_3 \times S_3 \times S_2 \times S_1. \end{aligned}$$

In general $S_{(\lambda_1,\lambda_2,\ldots,\lambda_k)}$ and $S_{\lambda_1} \times S_{\lambda_2} \times \cdots \times S_{\lambda_k}$ are isomorphic as groups.

Now $\pi \in S_n$ acts on a tableau $t = (t_{i,j})$ of shape $\lambda \vdash n$ as follows

$$\pi(t) = (\pi(t_{i,j}))$$

For example, if $\pi = (1, 2, 3)$,

$$\pi \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & \\ \cline{1-1} \end{array} = \begin{array}{|c|c|} \hline 2 & 3 \\ \hline 1 & \\ \cline{1-1} \end{array}$$

This induces an action on tabloids by letting

$$\pi\{t\} = \{\pi t\}.$$

This is well defined because it is independent of the choice of t.

Let X^λ denote the set of tabloids of shape λ. As we said, S_n acts transitively on X^λ by permuting the entries in the tabloids. In particular we observe that the Young subgroup S_λ is the subgroup of S_n stabilizing any given tabloid corresponding to λ.

DEFINITION 1.3.5 *The permutation representation associated to the action of S_n on X^λ is a vector space with basis*

$$\left\{e_{\{t\}}\right\}_{\{t\}\in X^\lambda}.$$

This space is denoted by M^λ and called the ***permutation module*** *corresponding to λ.*

As a representation

$$\begin{aligned}\rho_\lambda :& S_n \longrightarrow GL(M^\lambda)\\ &\pi \mapsto \rho_\lambda(\pi)\end{aligned}$$

where

$$\begin{aligned}\rho_\lambda(\pi) :& M^\lambda \longrightarrow M^\lambda\\ &e_{\{t\}} \mapsto e_{\pi\{t\}}\end{aligned}$$

This representation is reducible and contains all the irreducible representations of S_n we are looking for.

1.3.2 Specht modules

We now look for all the irreducible modules of S_n. These are the so–called Specht modules S^λ.

Any Young tableau naturally determines certain isomorphic copies of Young subgroups in S_n.

DEFINITION 1.3.6 *Let $\lambda = (\lambda_1, \ldots, \lambda_k) \vdash n$. Suppose that a tableau t has rows $R_1, R_2, \ldots, R_k$ and columns $C_1, C_2, \ldots, C_m$. Then*

$$R_t = S_{R_1} \times \cdots \times S_{R_k}$$

and

$$C_t = S_{C_1} \times \cdots \times S_{C_m}$$

are called the ***row–stabilizer*** *and the* ***column–stabilizer*** *of t, respectively.*

For each tableau t, define

$$e_t := \sum_{\pi \in C_t} \operatorname{sgn}(\pi) \cdot e_{\pi\{t\}}.$$

This is a linear combination of basis vectors of M^λ, thus $e_t \in M^\lambda$. We call such an element a **polytabloid**.

DEFINITION 1.3.7 *For any partition $\lambda \vdash n$, the corresponding **Specht module** S^λ is the subspace of M^λ spanned by the polytabloids $\{e_t\}$, where t is a tableau of shape λ.*

S^λ is an invariant subspace of M^λ under the action of S_n; indeed it is not difficult to check that $\sigma(e_t) = e_{\sigma(t)}$, for each $\sigma \in S_n$ and t tableau of shape λ; so for each $f \in S^\lambda$, then $\sigma(f) \in S^\lambda$.

THEOREM 1.3.1 (Submodule theorem, see Sagan [41] pag. 65) *Let U a submodule of M^λ. Then*

$$S^\lambda \subseteq U \quad \text{or} \quad U \subseteq {S^\lambda}^\perp.$$

In particular, when the field is $\mathbb{C}$, the S^λ are irreducible.

At this stage, we have one irreducible representation for each partition λ of n. As already observed, the number of irreducible representations of S_n equals to the number of conjugacy classes, which equals the number of partitions of n. Showing that all the S^λ are non–isomorphic, we prove that they are all the irreducible representations of S_n. This is proved by

THEOREM 1.3.2 (see Sagan [41] pag. 66) *For any $\lambda \vdash n$, the S^λ are all the irreducible representations of S_n.*

1.3.3 The decomposition of M^λ

The representation theory of S_n is studied by decomposing in a systematic way the permutation representation M^λ. Within each M^λ there is a uniquely determined irreducible subspace S^λ and letting λ running through all partitions of n accounts for all irreducible representations of S_n, without multiplicity.

Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ and $\mu = (\mu_1, \mu_2, \ldots, \mu_m)$ be two partitions of n. Then λ **dominates** μ, written $\lambda \trianglerighteq \mu$, if $\lambda_1 + \lambda_2 + \cdots \lambda_i \geq \mu_1 + \mu_2 + \cdots + \mu_i$, for all $i \geq 1$.

An easy corollary of theorem (1.3.2) is the following

THEOREM 1.3.3 (see Sagan [41] pag. 66) *Let $\mu \vdash n$ be a partition of n. The permutation module M^μ decomposes as*

$$M^\mu = \bigoplus_{\substack{\lambda \trianglerighteq \mu \\ \lambda \vdash n}} m_{\lambda\mu} S^\lambda \tag{1.5}$$

where $m_{\lambda\mu} := \dim \mathrm{Hom}(S^\lambda, M^\mu)$, and $m_{\mu\mu} = 1$.

We need to understand better how the M^λ decompose into irreducible subspaces. The coefficients $m_{\lambda\mu}$ have a nice combinatorial interpretation, which is explained in terms of semistandard tableaux.

DEFINITION 1.3.8 *Let $\lambda \vdash n$. A **generalized Young tableau** of shape λ is an array T obtained by replacing the nodes of λ with positive integers, repetitions allowed.*

We call **type** or **content** of T the composition $\mu = (\mu_1, \ldots, \mu_m)$, where μ_i equals the number of i's in T. Note that in a Young tableau (not generalized) entries may not be repeated. Let

$$T_{\lambda\mu} = \left\{ \begin{array}{l} T \text{ generalized Young tableau:} \\ T \text{ has shape } \lambda \text{ and content } \mu \end{array} \right\}.$$

DEFINITION 1.3.9 *A generalized tableau is said to be* ***semistandard*** *if its entries are nondecreasing across rows and increasing down columns.*

Let

$$T^0_{\lambda\mu} = \left\{ \begin{array}{l} T \text{ semistandard tableau:} \\ T \text{ has shape } \lambda \text{ and content } \mu \end{array} \right\}.$$

The Kostka numbers count semistandard tableaux.

DEFINITION 1.3.10 *Let $\lambda, \mu \vdash n$. The* ***Kostka numbers*** *are*

$$K_{\lambda\mu} = |T^0_{\lambda\mu}|$$

The following is well known

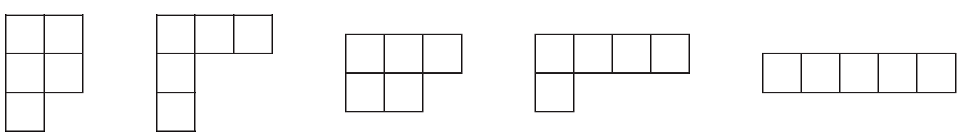
THEOREM 1.3.4 (Young's rule, see Sagan [41] pag. 85) *The multiplicity of S^λ in M^μ is equal to the number of semistandard tableaux of shape λ and content μ, i.e.*

$$M^\mu \cong \bigoplus_\lambda K_{\lambda\mu} S^\lambda \tag{1.6}$$

Is not difficult to see that we can restrict this direct sum to $\lambda \trianglerighteq \mu$.

For example, suppose that $\mu = (2,2,1)$. Then the possible $\lambda \trianglerighteq \mu$ and the associated semistandard tableaux of shape λ and content μ are

$(2,2,1)$ $\quad$ $(3,1,1)$ $\quad$ $(3,2)$ $\quad$ $(4,1)$ $\quad$ (5)



$(2,2,1)$:
$$\begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 & 2 \\ \hline 3 & \\ \hline \end{array}$$

$(3,1,1)$:
$$\begin{array}{|c|c|c|}\hline 1 & 1 & 2 \\ \hline 2 & & \\ \hline 3 & & \\ \hline \end{array}$$

$(3,2)$:
$$\begin{array}{|c|c|c|}\hline 1 & 1 & 2 \\ \hline 2 & 3 & \\ \hline \end{array} \qquad \begin{array}{|c|c|c|}\hline 1 & 1 & 3 \\ \hline 2 & 1 & \\ \hline \end{array}$$

$(4,1)$:
$$\begin{array}{|c|c|c|c|}\hline 1 & 1 & 2 & 2 \\ \hline 3 & & & \\ \hline \end{array} \qquad \begin{array}{|c|c|c|c|}\hline 1 & 1 & 2 & 3 \\ \hline 2 & & & \\ \hline \end{array}$$

(5):
$$\begin{array}{|c|c|c|c|c|}\hline 1 & 1 & 2 & 2 & 3 \\ \hline \end{array}$$

Thus

$$M^{(2,2,1)} \cong S^{(2,2,1)} \oplus 2S^{(3,3,1)} \oplus 2S^{(3,2)} \oplus S^{(4,1)} \oplus S^{(5)}.$$

Theorem (1.3.4) gives the isotypic decomposition of M^λ. It is a decomposition into (possibly) reducible components no one of which have any irreducible components in common. For the isotypic decomposition we can write

$$M^\mu = \bigoplus_\lambda I_{\lambda\mu} \tag{1.7}$$

where $I_{\lambda\mu}$ is the isotypic component corresponding to the partition λ, then

$$I_{\lambda\mu} \cong K_{\lambda\mu} S^\lambda. \tag{1.8}$$

Young's rule (see theorem (1.3.4)) tell us that the multiplicity of S^λ in M^μ equals the number of semistandard tableaux of shape λ and content μ.
A case in particular is very important for a peculiar class of examples of interest. Suppose that $\mu = (n-k, k)$, with $k \leq \frac{n}{2}$. We are looking for the decomposition of $M^{(n-k,k)}$; we need to know the possible shapes of λ allowed by Young's rule. They are

$$\begin{array}{ll} \underline{\overbrace{1\,1\ldots1}^{n-m}\;\overbrace{2\ldots2}^{m},\;\overbrace{1\,1\ldots1}^{n-m}\;\overbrace{2\ldots2}^{m-1},\ldots} & 1\,1\ldots1\,1\,1\ldots1 \\ \qquad\qquad\qquad\quad 2 & 2\,2\ldots2\,2 \end{array}$$

Each occurs once only. So the isotypic decomposition of $M^{(n-k,k)}$ (see [41], [18] and [11]) is

THEOREM 1.3.5 *The isotypic decomposition of $M^{(n-k,k)}$, with $k \leq \frac{n}{2}$, is given by*

$$M^{(n-k,k)} = S^{(n)} \oplus S^{(n-1,1)} \oplus \cdots \oplus S^{(n-k,k)}$$

where $S^{(n-j,j)}$ is an irreducible representation of S_n of dimension

$$\dim S^{(n-j,j)} = \binom{n}{j} - \binom{n}{j-1}.$$

We can translate this decomposition into an interpretation for what we call *the best of k out of n problem or data*; under this point of view the subspaces $S^{(n-k,k)}$ have this interpretations (this concept will be better explained in section 1.5 and Chapter 2 and 3):

$S^{(n)}$	corresponds to	the grand mean or the number of people in sample
$S^{(n-1,1)}$		the effect of item i, $1 \leq i \leq n$
$S^{(n-2,2)}$		the effect of items $\{i, j\}$, adjusted for the effect of i and j
$\vdots$		
$S^{(n-k,k)}$		the effect of a subset ok k items, adjusted for lower order effects.

The meaning of these statistical interpretation will be clearer in section 1.5 dedicated to spectral analysis.

1.4 Fourier transform on groups

The Fast Fourier Transform has a long and interesting history. Originally discovered by Gauss and later made famous by Cooley and Tukey (see [7]), it may be viewed as an algorithm which efficiently computes the Discrete Fourier Transform (DFT). Recently, there has developed a growing literature related to the construction of algorithms which generalize the FFT from the point of view of the theory of group representations (see [5], [11], [26] and [30]). These sort of generalizations are "natural" as mathematical constructs, but in point of fact they too have been motivated by applications, such as, as in our context, efficient data analysis (see Diaconis [9] and [10]).

DEFINITION 1.4.1 *Let $n \in \mathbb{N}$. The* ***DFT (Discrete Fourier Transform)*** *is a function $F : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ sending $(x_0, \ldots, x_{n-1})$ to $(X_0, \ldots, X_{n-1})$ where*

$$X_k := \sum_{j=0}^{n-1} x_j \, \omega^{jk} \tag{1.9}$$

with $k = 0, \ldots, n-1$ and $\omega = \mathrm{e}^{\frac{2\pi i}{n}}$.

We call **FFT (Fast Fourier Transform)** the family of algorithms which make the calculation of the DFT *fast*.

We observe that the DFT can be rewritten as a function $F : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, $F(f(0), \ldots, f(n-1)) = (\hat{f}(0), \ldots, \hat{f}(n-1))$, where $\hat{f}(k) = \sum_{j=0}^{n-1} f(j)\omega^{jk}$. Let $G = \mathbb{Z}/n\mathbb{Z}$, let χ_k be an irreducible character of G. Thus $\chi_k = \omega^{jk}$, so the DFT for the cyclic group of order n is

$$\hat{f}(k) = \sum_{j \in G} f(j)\chi_k(j), \tag{1.10}$$

that is a combination of the irreducible characters of G and the function f on G. The natural link between the classical definition of DFT and group representation theory is now clear. We can generalize to the following definition.

DEFINITION 1.4.2 *Let G be a finite group and $f : G \longrightarrow \mathbb{C}$ a complex–valued function on G. Let $\rho : G \longrightarrow GL(V)$ be a representation of G. Then the* ***Fourier transform of*** f ***at*** ρ *is*

$$\hat{f}(\rho) = \sum_{s \in G} f(s)\rho(s). \tag{1.11}$$

Similarly the ***Fourier transform of*** f ***at the matrix coefficient*** ρ_{ij} *is the scalar sum*

$$\hat{f}(\rho_{ij}) = \sum_{s \in G} f(s)\rho_{ij}(s).$$

A Fourier transform determines f through the Fourier inversion formula (see [10]).

PROPOSITION 1.4.1 (Fourier inversion formula) *Let G be a finite group, f a complex-valued function on G and $\mathcal{R}$ a set of irreducible representations of G. Then*

$$f(s) = \frac{1}{|G|} \sum_{\rho \in \mathcal{R}} d_\rho \operatorname{Trace}\left(\hat{f}\left(\rho\right) \rho\left(s^{-1}\right)\right)$$

where d_ρ is the degree of ρ.

In this generalized context the term FFT can be used in the same way to denote the collection of algorithms which make the calculation of the Fourier transform on finite groups efficient.

DEFINITION 1.4.3 *Let G be a finite group. Let $\mathcal{R}$ be a set of representations of G. We define the* ***complexity of the Fourier Transform for the set*** $\mathcal{R}$ *as*

$$\mathrm{T}_G(\mathcal{R}) := \min \# \text{ operations needed to compute the Fourier transform of } f \text{ on } \mathcal{R} \text{ via a straight-line program}$$

and the ***complexity of the group*** G *as*

$$\mathrm{C}(G) := \min \mathrm{T}_G(\mathcal{R})$$

where $\mathcal{R}$ varies over all complete sets of non–isomorphic irreducible representations of G.

We observe that a direct computation of the Fourier transform on a finite group requires $|G|^2$ operations, because a matrix–vector multiplication is involved. Many algorithms have been produced to improved this bound for certain classes of groups. We can summarize the current state of affair for some finite groups in the following table and list a large and growing bibliography (see [1], [2], [5], [11], [26], [27], [31], [38] and [39]),

G	C(G)	
abelian groups	$O(\|G\| \cdot \log \|G\|)$	(see [5])
symmetric groups	$O(\|G\| \cdot \log^2 \|G\|)$	(see [2])
wreath product of symmetric groups	$O(\|G\| \cdot \log^c \|G\|)$	(see [38])
supersolvable groups	$O(\|G\| \cdot \log \|G\|)$	(see [1])

It is conjectured that there is an $O(|G| \cdot \log^c |G|)$ upper-bound for all finite groups; this is one of the most important open problems in the field of Fourier transforms on finite groups.

1.4.1 Fourier transform on the symmetric group

The computation of Fourier transform on symmetric groups was first studied by Clausen (see [5]), then by Diaconis and Rockmore (see [11]) and Maslen (see [27]).

THEOREM 1.4.1 (see Clausen, [5]) *Let S_n be the symmetric group. Then*

$$C(S_n) < \frac{1}{2}(n^3 + n^2)n!$$

Since $\log(n!)$ is of order $n \cdot \log n$, this result can be reformulated as

$$C(S_n) = O(|S_n| \cdot \log^3 |S_n|).$$

Diaconis and Rockmore (see [11]) obtained this general result:

THEOREM 1.4.2 (see Diaconis and Rockmore, [11])
Let S_n be the symmetric group. Let $T(n)$ be the number of operations required to compute the Fourier transform of a function on S_n at all irreducible representations. Then

$$T(n) \leq B \cdot (n!)^{\frac{a}{2}} \cdot n \cdot e^{-(a-2)c\sqrt{\frac{n}{2}}}$$

where B is a positive computable constant, $a > 2$ is the exponent for matrix multiplication (i.e. multiplying $d \times d$ takes d^a operations), $c = 0.1156$.

Currently the best theoretical result for a is $a = 2.38$ (see [8]).

Maslen (see [27]) obtained a refinement of Clausen's bound; he replaced the matrix multiplications in Clausen's algorithm with sums indexed by combinatorial objects that generalize Young tableaux.

THEOREM 1.4.3 (see Maslen, [27]) *Let f be a complex function on S_n. Then the Fourier transform of f at a complete set of irreducible representations in Young's orthogonal form may be computed in no more than*

$$\frac{3n(n-1)}{4}|S_n|$$

multiplications and the same number of additions.

This result can be reformulated as

$$C(S_n) = O(|S_n| \cdot \log^2 |S_n|).$$

1.5 Spectral analysis on groups

Spectral analysis is a non–model based approach to data analysis, formulated in general group theoretic setting by Diaconis (see [9] and [10]); it extends the classical spectral analysis of time series and for this reason it is called also **generalized spectral analysis**. Often data are presented as a function $f(x)$ defined on some index set X. If X is connected to a group, the function f can be *Fourier expanded* and one may try to interpret its coefficients. More precisely, the idea of spectral analysis is that often data has natural symmetries, encapsulated in the existence of a symmetry group for the domain of the data. The organizing principle of spectral analysis is the understanding of data trough its decomposition according to these symmetries.

We recall from section 1.2 that if X is a finite set, G a group acting on X and $L(X)$ the vector space of complex–valued functions on G, then $L(X)$ may be decomposed as an orthogonal direct sum of G-invariant subspaces

$$L(X) = V_1 \oplus \cdots \oplus V_h, \tag{1.12}$$

called the isotypic decomposition. Spectral analysis takes the form of computing the projections of the data onto these subspaces and judging which projections are significant. The effect is to represent the data vector f as the sum

$$f = f_1 + \cdots + f_h, \tag{1.13}$$

where f_i is the projection of f on V_i. The **normalized length** of a given projection indicates its influence on the data; a large projection may suggest that further investigation is merited.

In general the theme is that under the assumption of a natural symmetry group for the domain of the data, group theory can be used to decompose the data as well as to indicate expansions of the data which will help in its interpretations. The use of generalized FFTs for spectral analysis is the efficient computation of the projections.

1.5.1 Time series analysis

This very general principle encompasses various standard approach to data analysis. A widely studied example is the **analysis of time series** (see [6] and [36]).

In this situation the goal is to analyze some function of time, say the Dow Jones average, seismograph data, or the number of babies born in Rome each day, by expanding the observed function into sum of sines and cosines. The expansion obtained is precisely the Fourier expansion and analysis proceeds by looking for the large Fourier coefficients, i.e. the large projections. Computation requires a discretization and truncation of the data and in so doing the expansion is computed as a discrete Fourier transform and it is performed efficiently by the abelian Fast Fourier Transform.

Example 1: Time series analysis.

Diaconis (see [9]) provides a simple but clarifier example. Suppose to have the data on the number of babies born daily in New York City over a five year period. Here $X = \{1, \dots, n\}$ where $n = 365 \times 5 + 1$. The data are represented as a function

$$f(x) = \text{ number of born on day } x.$$

Izenman and Zabell carried out these studies in 1978 (see [17]); inspecting the data given by the function $f(x)$ they found strong periodic phenomena: about 450 babies were born on each week day and about 350 on each day of the weekend. There might be also monthly and quarterly effects.

To examine such a phenomena, we may pass from the original data $f(x)$ to its Fourier transform

$$\hat{f}(y) = \sum_{x=0}^{n-1} f(x) e^{\frac{2\pi i x y}{n}}.$$

Fourier inversion formula (see proposition (1.4.1); here it is in its classical abelian version) gives

$$f(x) = \frac{1}{n} \sum_{y=0}^{n-1} \hat{f}(y) e^{-\frac{2\pi i x y}{n}}.$$

It sometimes happens that a few values of $\hat{f}(y)$ are much larger than the rest and determine f in the sense that f is closely approximated by the function defined by using only the large Fourier coefficients in the inversion formula. When this happens, we have f approximated by few simple periodic functions on x, $e^{-\frac{2\pi ixy}{n}}$, and may feel to understand the situation.
The hunting and interpretation of periodicities is one use of spectral analysis.

1.5.2 Spectral analysis of full and partially ranked data

Data sometimes come in the form of rank of preference. For example, elections are sometimes based on ranking (this happens in some Australian elections, for instance).

Most anyone who analyzes such data looks at simple averages, such as the proportion of times each item was ranked first (or last) and the average rank for each item. There are **first order statistics**: they are linear combinations of the number of times item i was ranked in position j. There are also natural **second order statistics**, based on the number of times items i and j are ranked in position k and l; these come in ordered and unordered modes: for example, the number of times items i and j are ranked either 12 or 21 is an unordered second order statistic. Similarly there are third and **higher order statistics** of various type.

Diaconis in [9] underlines this crucial point: "*a basic idea of data analysis is this: if you've found some structure, take it out and look at what is left. Thus, to look at second order statistic it is natural to substract away the observed first order structure. This leads to a natural decomposition of the original data into orthogonal pieces*".

Suppose that individuals are given a list of items and asked to rank them or some subset of them in terms of preference. The requested ranking may be *full*, in the sense that the respondent is asked to reorder the entire list, or *partial*, meaning that only a subset is to be chosen for ranking.

Example 2: Full rankings.

Diaconis (see [9]) illustrates how the *Fourier analysis* may work in a *full ranking* problem.
Suppose that people are asked to rank where they want to live: in a city, suburbs or country. They are asked to rank these three items. Suppose the rankings are

π	city	suburbs	country	#
id	1	2	3	242
(23)	1	3	2	28
(12)	2	1	2	170
(132)	3	1	2	628
(123)	2	3	1	12
(13)	3	2	1	359

Here $X = S_3$ and

$$f(x) = \text{ number of people choosing } \pi.$$

In order to have the *Fourier expansion* of f, we need to know the irreducible representations of S_3. They are the trivial, the sign-representation and a two–dimensional representation ρ. The Fourier inversion formula (see proposition (1.4.1)) gives

$$f(\pi) = \frac{1}{6}\left[\hat{f}(\mathrm{triv}) + \mathrm{sgn}(\pi)\hat{f}(\mathrm{sgn}) + 2\,\mathrm{Tr}(\rho(\pi^{-1})\hat{f}(\rho))\right].$$

Expanding the trace gives a spectral analysis of f as a sum of orthogonal functions.
To facilitate comparisons between functions in this basis, let us choose an orthogonal version of ρ. Thus

π	id	(12)	(13)
$\rho(\pi)$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$

π	(23)	(123)	(132)
$\rho(\pi)$	$\frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}$

They are arrived choosing $w_1 = \frac{1}{2}(e_1 - e_2)$ and $w_2 = \frac{1}{\sqrt{6}}(e_1 + e_2 - 2e_3)$ as an orthogonal basis for $\{v \in \mathbb{R}^3 : v_1 + v_2 + v_3 = 0\}$. The matrices $\rho(\pi)$ give the action of π in this basis.

Now

$$\begin{aligned}
\hat{f}(\text{triv}) &= 1439 \\
\hat{f}(\text{sgn}) &= 242 - 28 - 170 + 628 + 12 - 359 = 325 \\
\hat{f}(\rho) &= \begin{pmatrix} -54.5 & \frac{285\sqrt{3}}{2} \\ \frac{-947\sqrt{3}}{2} & -101.5 \end{pmatrix}
\end{aligned}$$

Define four functions on S_3 by

$$\sqrt{2}\rho(\pi^{-1}) = \begin{pmatrix} a(\pi) & b(\pi) \\ c(\pi) & d(\pi) \end{pmatrix}$$

With this definition, the functions id, $\text{sgn}(\pi)$, $a(\pi)$, $b(\pi)$, $c(\pi)$, $d(\pi)$ are orthogonal and have the same length.
Expanding the trace in Fourier inversion formula gives

$$\begin{aligned}
f(\pi) = \frac{1}{6}\Bigg[&1429 + 325\,\text{sgn}(\pi) - 54.5\sqrt{2}a(\pi) - 947\sqrt{\frac{3}{2}}b(\pi) + \\
&+285\sqrt{\frac{3}{2}}c(\pi) - 101.5\sqrt{2}d(\pi)\Bigg] = \\
&= \frac{1}{6}\left[1439 + 325\,\text{sgn}(\pi) - 77a(\pi) - 1160b(\pi) + 349c(\pi) - 144d(\pi)\right].
\end{aligned}$$

As a check, when $\pi = \mathrm{id}$, this becomes $242 = \frac{1}{6}[1439 + 325 - 109 - 203]$.

The largest non–constant coefficient is 1160 and multiplies $b(\pi)$. This is the function

π	id	(12)	(13)	(23)	(123)	(132)
$b(\pi)$	0	0	$\sqrt{\frac{3}{2}}$	$-\sqrt{\frac{3}{2}}$	$\sqrt{\frac{3}{2}}$	$-\sqrt{\frac{3}{2}}$

or

$$b(\pi) = \begin{cases} -\sqrt{\frac{3}{2}} & \text{if cities are ranked 3rd } (\pi(1) = 3) \\ 0 & \text{if country is ranked 3rd } (\pi(3) = 3) \\ \sqrt{\frac{3}{2}} & \text{if suburbs are ranked 3rd } (\pi(2) = 3) \end{cases}$$

Spectral analysis gives fresh insight into thus little data set: after the constant, the best *single predictor of f is what people rank last.*
Now $b(\pi)$ enters with a negative coefficient. This means that people *hate* the city most, the suburbs least and the country in between. Going back to the data,

$$\begin{aligned} \#\{\pi(1) = 3\} &= 981 \\ \#\{\pi(2) = 3\} &= 40 \\ \#\{\pi(3) = 3\} &= 412 \end{aligned}$$

so the effect is real.

Example 3: Full rankings (again).

Rockmore in [39] proposes another very enlightening example. A movie studio is interested in current viewing trends. Respondent are presented with the list of movie

1. The Lord of the Rings: The Return of the King
2. Spiderman II
3. Lost in translation
4. Pulp Fiction
5. Hidalgo

and asked to rank them in order of preference. A possible response may be 1-2-5-3-4, equivalent to the choice of permutation 12534. If many people are asked, then a function

$$\begin{aligned} f &: S_5 \longrightarrow \mathbb{Z} \\ f(\pi) &= \text{ number of respondent with preference order } \pi \end{aligned}$$

is determined.

Let X be the set of items to be ranked, $G = S_5$ and $L(X)$ the vector space of all complex–valued functions on S_5. $L(X)$ decomposes into the direct sum of seven subspaces

$$L(X) = V_1 \oplus V_2 \oplus \cdots \oplus V_7 \tag{1.14}$$

where the dimensions of the V_i's are

	V_1	V_2	V_3	V_4	V_5	V_6	V_7
$\dim V_i$	1	16	25	36	25	16	1

There are **natural statistics** to compute from data $f(\pi)$.

The first thing to be computed is the **mean**, or average response. This is precisely the projection of the data onto V_1, the one-dimensional *space of* *constant functions.* It is given by the constant vector all of whose entries are equal to

$$\frac{1}{|S_5|}\sum_{\pi\in S_5} f(\pi).$$

Restarted in terms of representation theory, this is essential the Fourier transform of the data at the trivial representation.

Next a **first order summary** of the data is obtained by counting how many respondents ranked movie i in position j. Notice that this is precisely the content of the Fourier transform of the data at the defining representation of S_5. That is, define

$$\begin{aligned} \rho_{ij} &: S_5 \longrightarrow \mathbb{Z} \\ \rho_{ij}(\pi) &:= \begin{cases} 1 & \text{if } \pi(i) = j \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Then the Fourier transform of f at ρ_{ij} is

$$\begin{aligned} \hat{f}(\rho_{ij}) &= \sum_{\pi\in S_5} f(\pi)\rho_{ij}(\pi) \\ &= \text{ the number of respondent ranking movie } i \text{ at position } j \end{aligned}$$

The term $\rho_{ij}(\pi)$ is the (i,j)-entry of the matrix corresponding to the representation $\rho(\pi)$, which is the representation of S_5 assigning to π the corresponding "permutation matrix". The first order analysis consists of computing the Fourier transform $\hat{f}(\rho)$.
In this case we have the projection of the data on V_2, which is called the *space of first order functions.* A general first order function has the form

$$\sum_{i,j=1}^{5} a_{ij}\rho_{ij}(\pi)$$

where the coefficients a_{ij} must satisfy $\sum_{i,j=1}^{5} a_{ij} = 0$ (to get a direct sum decomposition).

Similarly **higher order summaries** can be obtained by computing Fourier transforms at other representations of S_5. A higher order effect attempt to account for interactions in the data. Relevant functions for these higher order projections would be

$$\rho_{\{i,j\},\{k,l\}} : S_5 \longrightarrow \mathbb{Z}$$
$$\rho_{\{i,j\},\{k,l\}}(\pi) := \begin{cases} 1 & \text{if } \pi(\{i,j\}) = \{k,l\} \\ 0 & \text{otherwise} \end{cases}$$

or their "ordered version"

$$\rho_{(i,j),(k,l)} : S_5 \longrightarrow \mathbb{Z}$$
$$\rho_{(i,j),(k,l)}(\pi) := \begin{cases} 1 & \text{if } \pi((i,j)) = (k,l) \\ 0 & \text{otherwise} \end{cases}$$

In this case we have the projection of the data on V_3, the *space of unordered second order functions* and on V_4, the *space of ordered second order functions*.
A typical unordered second order function may be

$$\sum_{i,j,k,l=1}^{5} a_{ijkl}\rho_{\{i,j\},\{k,l\}}(\pi)$$

with a_{ijkl} chosen such that V_3 is orthogonal to $V_1 \oplus V_2$.
A typical ordered second order function may be

$$\sum_{i,j,k,l=1}^{5} a_{ijkl}\rho_{(i,j),(k,l)}(\pi)$$

with a_{ijkl} chosen such that V_4 is orthogonal to $V_1 \oplus V_2 \oplus V_3$.
$\hat{f}(\rho_{\{i,j\},\{k,l\}})$ records the number of people ranking movies i and j in position k and l, where order is not important, while in $\hat{f}(\rho_{(i,j),(k,l)})$ the order is important.
Implicit here is the computation of Fourier transforms for functions on the symmetric group at well–known reducible representations given by actions on Young tableaux.

1.5.3 General considerations

As we shown in section 1.3, the representation theory of S_n is studied by decomposing in a systematic way the permutation module M^λ.
The discussion in Example 3 shows that the Fourier transforms at matrix coefficients of the representations on the reducible module M^λ are easily interpreted.
However, a spectral analysis representation rewrites the function in terms of Fourier transforms at irreducible matrix coefficients. The claim is that the Fourier trasforms at the matrix coefficients for a given S^λ encode the pure interactions specified by λ.

Consider, for example, the representation $M^{(n-1,1)}$ given by the symmetric group S_n acting on Young tableaux of shape $(n-1,1)$.
Any such Young tableau is determined by the entry in the second row, and thus, may be identified with the action of S_n on the standard basis $e_1,\dots,e_n$ given by $\rho(\pi)(e_i)=\rho_{\pi(i)}$, which is the defining representation of S_n.
For any fixed i, the matrix coefficients $\{\rho_{i1},\dots,\rho_{in}\}$ span an S_n-invariant subspace of $L(S_n)$. This follows from the fact that $\pi(\rho_{ij}(\sigma))=\rho_{ij}(\pi^{-1}\sigma)$, implying that $\pi(\rho_{ij})=\rho_{i,\pi(j)}$.
Thus, the set of matrix coefficients $\{\rho_{i1},\dots,\rho_{in}\}$ do themselves span a copy of $M^{(n-1,1)}$, thereby providing n easily identified isomorphic copies of the space $M^{(n-1,1)}$.
According to theorem (1.3.4), the representation space $M^{(n-1,1)}$ decomposes as

$$M^{(n-1,1)}=S^{(n)}\oplus S^{(n-1,1)}$$

where $S^{(n)}$ denotes the trivial representation, spanned by the subspace of vectors with constant coordinates, while $S^{(n-1,1)}$ denotes its $(n-1)$-dimensional orthogonal irreducible complement of those vectors whose coordinates sum to zero. These copies of $M^{(n-1,1)}$ are mutually orthogonal. For example, each contains the same copy of the trivial representation.
Also, notice that any one of the first order statistics $\{\hat{f}(\rho_{i1}),\dots\hat{f}(\rho_{in})\}$ is determined by knowing all the others. In fact, the matrix coefficients span a space of dimension $((n-1)^2+1)$.
Herein lies the connection between the Fourier transform $\hat{f}(\rho)$ and the decomposition of the original data vector.

The entire vector space $L(S_n)$ has the isotypic decomposition

$$L(S_n) = \bigoplus_{\lambda \vdash n} I^{\lambda} \tag{1.15}$$

where each subspace I^{λ} is equivalent to d_{λ} copies of S^{λ}.
The space of matrix coefficients ρ_{ij} span a subspace of $L(S_n)$ that has an S_n-irreducible decomposition isomorphic to

$$S^{(n)} \oplus S^{(n-1,1)}$$

and the computation of $\hat{f}(\rho)$ is equivalent to computing the projection of f onto the trivial representation, as well as the isotypic component of $L(S_n)$ which corresponds to the irreducible representation $S^{(n-1,1)}$, denoted as $I^{(n-1,1)}$ in (1.15). The projections are the Fourier transforms at the corresponding irreducible representations and in this case, the projection onto $I^{(n-1,1)}$ encodes the first order information about f.

A similar argument holds true for higher order statistics as well.

Thus, the summary is that, for each partition λ fo n, there is a permutation representation M^{λ}. The matrix coefficients do themselves give a representation and the Fourier transform of the data computes the projection of the data onto this invariant subspaces. In the natural basis of the represenation, the corresponding Fourier transform at this basis computes certain frequency counts, but this information is both coarse and redundant. Obtaining the **pure higher order effect** (as represented by λ) is equivalent to the computation of the projection of the data onto the S^{λ}-isotypic, which is the same of computing the Fourier transform of the data at the irreducible representation corresponding to λ.

Example 4: Partial rankings.

Rockmore in [39] re-discusses from a spectral theory point of view the data obtained by Thompson (see [43]) from the Catholic Charities Organitazion.

Catholic Charities sent out a questionairre to a sample of its members, asking each participant to choose in order three of eleven possible charitable directions, that is an ordered 3–set from a 11–set.

To analyze such a study, many questions arise. Are people mainly choosing a favorite or two and then "randomly" choosing the rest? These would be first order considerations. Or are people's choices driven by commonalities among the options? These are higher order effects.

We will explain in some part the spectral analysis approach to this data, which seeks to analyze the data as element of $M^{(8,1,1,1)}$, a representation space of S_{11}.

Let $\mu = (n-1,1)$. Then the Kostka numbers are $K_{\lambda\mu} = h(\lambda) - 1$, so, according to Young's rule (see theorem (1.3.4),

$$I_{\lambda,(n-1,1)} \cong (h(\lambda)-1)S^{(n-1,1)}.$$

The first row of the Young diagram of shape $(n-1,1)$ can contain any string of nondecreasing entries, so that the only nonpermissible entry in the second row is 1. Thus, the various semistandard tableaux are determined by the entry in the second row which can be among $2,\ldots,h(\lambda)$.
For example, for $\mu = (8,1)$ and $\lambda = (4,2,1,1)$, we have $K_{\lambda\mu} = 3$, indeed

$$\begin{array}{|c|c|c|c|c|c|c|c|}\hline 1&1&1&1&2&2&3&4\\ \hline 2\\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|c|c|c|c|c|c|}\hline 1&1&1&1&2&2&2&4\\ \hline 3\\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|c|c|c|c|c|c|}\hline 1&1&1&1&2&2&2&3\\ \hline 4\\ \cline{1-1}\end{array}.$$

So we have 3 copies of $S^{(8,1)}$ in $M^{(4,2,1,1)}$.

It is useful to give an illustration of the $(n-1,1)$ isotypic in terms of relevant characteristics of partially ranked data.
Partially ranked data of shape $\lambda = (\lambda_1,\ldots,\lambda_k)$ can be viewed as frequencies of respondents picking out their favorite λ_k items, then their favorite λ_{k-1} items, and so on, all the way up to their least favorite λ items. Note that such a choice is determined by making only the first $k-1$ sets of choices.
Consider the following functions

$$\Delta_i^{(j)}(t) = \begin{cases} 1 & \text{if } t \text{ ranks } i \text{ among the } j^{th} \text{ favorites} \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\Delta_i^{(j)} = \frac{1}{c_{ij}} \sum \delta_t$$

where c_{ij} is an appropriate normalization constant and the sum runs over all tableaux t such that row $(k-j+1)$ contains i.

It is not difficult to see that the functions $\{\Delta_1^{(j)}, \ldots, \Delta_n^{(j)}\}$ span a subspace isomorphic to $M^{(n-1,1)}$. Following theorem (1.3.4), we have

$$M^{(n-1,1)} \cong S^{(n)} \oplus S^{(n-1,1)}.$$

$S^{(n)}$ is the subspace of constant functions and $S^{(n-1,1)}$ is its orthogonal complement consisting of those functions whose values sum to 0.
Letting j vary from 1 to $(n-1)$, we construct $(n-1)$ subspaces which only intersect in the one–dimensional subspace of constant functions. These various subspaces of individual ranked popularity are naturally viewed as subspaces of the first order effects.

Data analysis. Spectral analysis may in general proceed into two steps:

(1) first of all the coarse decomposition of the data vector f into its isotypic components is considered

Having done (1), the **lengths** of the projection are considered. If a given projection has a large relative contribution then

(2) it is further investigated by considering some irreducible decomposition of particular isotypic.

We are interested in the analysis of data of the shape $(8,1,1,1)$. Let f denote the original data vector. Theorem (1.3.4) gives us the decomposition of $M^{(8,1,1,1)}$ in

$$M^{(8,1,1,1)} = S^{(11)} \oplus 3S^{(10,1)} \oplus 3S^{(9,2)} \oplus 3S^{(9,1,1)} \oplus S^{(8,3)} \oplus 2S^{(8,2,1)} \oplus S^{(8,1,1,1)}$$

Analyzing data of the Catholic Charities, Rockmore obtained table (1.1) (see Rockmore, [39])
The isotypic decomposition indicates where the interesting projections are.

The large size of the $(10,1)$ projections suggests that this projection merits further analysis. We recall that this space measures the individual effects of the attraction (or repulsion) of individual charities.

To proceeds Rockmore follows the so–called *Mallow's method*, which considers inner products of the projection with naturally interpretable functions

μ	(11)	(10,1)	(9,2)	(9,1,1)	(8,3)	(8,2,1)	(8,1,1,1)
d_μ	1	10	44	45	110	231	120
$\dim(I_{\lambda,\mu})$	1	30	132	135	110	462	120
$\|\|f^\mu\|\|^2$	335.127	597.894	556.519	137.875	220.547	316.825	59.212

Table 1.1: Results

in the isotypics of interest (this method will be largely explained in Chapter 2. For references see [25]).

For the $(10,1)$–isotypic one natural set of spanning functions are the $\Delta_i^{(j)}$ defined by

$$\Delta_i^{(j)}(t) = \begin{cases} 1 & \text{if } i \text{ is ranked in position } j \\ 0 & \text{otherwise} \end{cases}$$

The following table gives the inner products of the $(10,1)$–isotypic projection with $\{\Delta_i^{(j)} : i = 1,2,3, j = 1,\ldots,11\}$

j / i	1	2	3	4	5	6	7	8	9	10	11
1	-37.4	-10.4	5.6	149.6	0.6	-12.4	-51.4	-6.4	-45.4	3.6	3.6
2	-28.4	9.6	28.6	59.6	10.6	-8.4	-43.4	-16.4	-42.4	43.4	-13.4
3	-34.4	22.6	37.6	8.6	7.6	-5.4	-31.4	-10.4	-31.4	28.6	7.6

Table 1.2: Results of Mallow's method

A quick inspection of the entries of table (1.2) shows a very large $(1,4)$ entry, indicating that many respondents fell strongly about choosing charity 4 first. If we refer back to table (1.1), we notice that the counts for 3–tuples which first entry 4 are by and large the greatest.
Entries $(3,2)$, $(2,3)$, $(3,3)$, $(2,4)$, $(2,10)$, $(3,10)$ are of the next scale, indicating strong interest in charities 2, 3 and 10.

2
Spectral analysis and voting

Our work, exposed in Chapter 3, is an application of generalized spectral analysis to economic science and in particular to preference formation.
It is modeled on some recent works of M. E. Orrison and B. L. Lawson (see [22], [23] and [24]) on noncommutative harmonic analysis of political voting. In this Chapter we want to illustrate this political application.

2.1 Noncommutative harmonic analysis of voting in committees

The kind of spectral analysis analyzed in section 1.5 of Chapter 1 is called generalized spectral analysis or also **noncommutative harmonic analysis** because it is a generalization of classical spectral analysis. Spectral analysis is the discrete Fourier analysis and it is basic for time series analysis and other types of analysis in the computational science, engineering and natural sciences. Diaconis (see [9] and [10]) extended the classical spectral analysis of time series to a non–time series subject, for the analysis of discrete data which has a noncommutative structure.

New efforts have been made in order to apply spectral analysis to a non–time series subject in the political sciences, above all in the analysis of voting.

Spectral analysis has been already used in political science to identify cycles in time series data. For example, spectral analysis was used to test if presidential popularity in the United State and the concentration of international power have periodic components.

Recently, M. E. Orrison and B. L. Lawson (see [22] and also [23], [24] with David T. Uminsky) introduced a generalization of spectral analysis as a new instrument for political scientist; they used the powerful machinery of spectral analysis to analyze political voting data. In particular, they analyzed votes of the nine judges of the United States Supreme Court (Warren Court 1958– 1962, Burger Court 1967–1981, Renquist Court 1994–1998) and detected influential coalitions.

With this theory political scientist can use spectral analysis as a method for identifying substantively important dynamics in politics, rather then just as a diagnostic tool.

The idea followed by Orrison and Lawson is to consider political voting data as elements of a mathematical framework; then the features of that framework can be used to work out natural interpretations of the data. The mathematical framework corresponding to voting data has many components, each of which encapsulates information on particular *coalition effects*; the decomposition of data with respect to these components provides the identification of **influential coalitions**.

2.2 Coalitions

Let $X = \{x_1, \ldots, x_n\}$ be a finite set and $f : X \longrightarrow \mathbb{C}$ a complex–valued function on X. Let M be the vector space of all complex–valued functions on X and S_n the symmetric group of order n.
As already explained in section 1.2, M may always be decomposed into a direct sum

$$M = M_0 \oplus \cdots \oplus M_h \tag{2.1}$$

for some positive integer h, where each M_i is an invariant subspace of M.

In particular, each function $f \in M$ may be written uniquely as a sum

$$f = f_0 + \cdots + f_h \tag{2.2}$$

with $f_i \in M_i$ and $\pi(f_i) \in M_i$, for all $\pi \in S$.

There are many ways to decompose M as the direct sum of invariant subspaces; the idea behind spectral analysis is to choose the decomposition of M that provides invariant subspaces that encapsulate important properties of the data.

Suppose that $X = \{X_1, \ldots, X_n\}$ is a set of n voters. Assume we have the results of N non–unanimous votes and that each person casts a ballot on each vote. They define

$$X^{(n-k,k)} = \text{ the set of } k\text{-elements subsets of the voters of } X \tag{2.3}$$

with $1 \leq k \leq \frac{n}{2}$ and denote with $f^{(n-k,k)}$ a function on $X^{(n-k,k)}$ defined as

$$f^{(n-k,k)}(\omega) = \text{ the number of times that } \omega \text{ is in the minority} \tag{2.4}$$

for each $\omega \in X^{(n-k,k)}$. Define

$$M^{(n-k,k)} = \text{ the vector space of all complex–valued functions on } X^{(n-k,k)}.$$

We observed that the permutations of S_n act on X, but also on the subsets in $X^{(n-k,k)}$, for each k. Then, as outlined in equation (2.1), $M^{(n-k,k)}$ may be decomposed as a direct sum

$$M^{(n-k,k)} = M_0 \oplus M_1 \oplus \cdots \oplus M_k \tag{2.5}$$

where each M_i is a subspace of $M^{(n-k,k)}$ invariant with respect to the action of S_n. The space M_0 is said to be corresponding to the **mean response**, that is the average number of times an element of $M^{(n-k,k)}$ is in the minority. M_1 corresponds to the so–called **first order effects**, whereas M_i is related to higher order effects, called **coalition effects**.

Spectral analysis focuses on the computation of the decomposition of each function $f \in M^{(n-k,k)}$ onto the components of (2.5), that is

$$f = f_0 + \cdots + f_k. \tag{2.6}$$

2.3 A five people committee

Let $X = \{A, B, C, D, E\}$ be a committee of five people and suppose we have the results of 128 non–unanimous votes. Data is viewed as a function f defined on the subsets of X; in particular we have

$$f^{(4,1)} = \text{the number of times one person of } X \text{ is in the minority.}$$

$$f^{(3,2)} = \text{the number of times two people of } X \text{ are in the minority.}$$

Suppose that

$$f^{(4,1)} = \begin{pmatrix} 10 \\ 9 \\ 3 \\ 2 \\ 1 \end{pmatrix} \begin{matrix} A \\ B \\ C \\ D \\ E \end{matrix} \qquad \text{and} \qquad f^{(3,2)} = \begin{pmatrix} 22 \\ 21 \\ 24 \\ 11 \\ 5 \\ 2 \\ 10 \\ 2 \\ 1 \\ 5 \end{pmatrix} \begin{matrix} AB \\ AC \\ AD \\ AE \\ BC \\ BD \\ BE \\ CD \\ CE \\ DE \end{matrix}$$

This means that, in this example, A is in the minority against the other four people for 10 times, whereas AB are in the minority against the other three for 22 times, and so on.

Let M be the vector space of the complex–valued functions on $X^{(4,1)}$ and $X^{(3,2)}$; M may be naturally decomposed as

$$M = M^{(4,1)} \oplus M^{(3,2)},$$

where $M^{(4,1)}$ is the subspace of the functions on $X^{(4,1)}$ and $M^{(3,2)}$ on $X^{(3,2)}$. These two subspaces may be again decomposed into invariant subspaces

$$M^{(4,1)} = M_0^{(4,1)} \oplus M_1^{(4,1)} \tag{2.7}$$

$$M^{(3,2)} = M_0^{(3,2)} \oplus M_1^{(3,2)} \oplus M_2^{(3,2)}. \tag{2.8}$$

We may project the functions $f^{(4,1)}$ and $f^{(3,2)}$ onto these invariant subspaces and obtain

$$f^{(4,1)} = f_0^{(4,1)} + f_1^{(4,1)} \tag{2.9}$$

$$f^{(3,2)} = f_0^{(3,2)} + f_1^{(3,2)} + f_2^{(3,2)}. \tag{2.10}$$

2.3.1 One person in the minority.

Going back to the example, according to decomposition (2.9), we get

$$f^{(4,1)} = \begin{pmatrix} 10 \\ 9 \\ 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \\ 5 \\ 5 \\ 5 \end{pmatrix} + \begin{pmatrix} 5 \\ 4 \\ -2 \\ -3 \\ -4 \end{pmatrix} \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array}$$

where

$$f_0^{(4,1)} = (5, 5, 5, 5, 5)$$

$$f_1^{(4,1)} = (5, 4, -2, -3, -4)$$

The number of votes in which one person is in the minority is 25, so the average of the individual minority is $5 = 25/5$; then $f_0^{(4,1)}$ is the mean response function. The function $f_1^{(4,1)}$ shows the **first order effect**, which counts the number of votes in which each person differs from the mean. In this case the interpretation of the first order effects doesn't yield new information in relation to the initial data; the largest value is for A, that is most often in the minority, and the smallest value is for E, that is less often in the minority.

2.3.2 Two people in the minority.

We can appreciate the power of spectral analysis in the analysis of higher order effects. According to decomposition (2.10), we obtain

$$f^{(3,2)} = \begin{pmatrix} 22 \\ 21 \\ 24 \\ 11 \\ 5 \\ 2 \\ 10 \\ 2 \\ 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 10.3 \\ 10.3 \\ 10.3 \\ 10.3 \\ 10.3 \\ 10.3 \\ 10.3 \\ 10.3 \\ 10.3 \\ 10.3 \end{pmatrix} + \begin{pmatrix} 11.53 \\ 8.20 \\ 9.53 \\ 7.53 \\ -4.80 \\ -3.47 \\ -5.47 \\ -6.80 \\ -8.80 \\ -7.47 \end{pmatrix} + \begin{pmatrix} 0.17 \\ 2.50 \\ 4.17 \\ -6.83 \\ -0.50 \\ -4.83 \\ 5.17 \\ -1.50 \\ -0.50 \\ 2.17 \end{pmatrix} \begin{matrix} AB \\ AC \\ AD \\ AE \\ BC \\ BD \\ BE \\ CD \\ CE \\ DE \end{matrix}$$

where

$$\begin{aligned} f_0^{(3,2)} &= (10.3, 10.3, 10.3, 10.3, 10.3, 10.3, 10.3, 10.3, 10.3, 10.3) \\ f_1^{(3,2)} &= (11.53, 8.20, 9.53, 7.53, -4.80, -3.47, -5.47, -6.80, -8.80, -7.47) \\ f_2^{(3,2)} &= (0.17, 2.50, 4.17, -6.83, -0.50, -4.83, 5.17, -1.50, -0.50, 2.17) \end{aligned}$$

The function $f_0^{(3,2)}$ is the mean response function; the number of votes in which two people are in the minority is 103, then the average of the minority of pairs is $10.3 = 103/10$. The functions $f_1^{(3,2)}$ and $f_2^{(3,2)}$ capture the **first order and second order effects**. In order to interpret these effects, Orrison and Lawson [22] suggest to use *Mallow's method* (see [25]).

2.3.3 Interpretation: Mallow's method

To interpret the first order effects, for each subset of voters H, define a function $f_H \in M^{(3,2)}$ which identifies the elements of $f^{(3,2)}$ "containing" H with 1 and those "not containing" H with 0. In particular,

$$\begin{aligned}
f_A &= (1,1,1,1,0,0,0,0,0,0)\\
f_B &= (1,0,0,0,1,1,1,0,0,0)\\
f_C &= (0,1,0,0,1,0,0,1,1,0)\\
f_D &= (0,0,1,0,1,0,0,1,0,1)\\
f_E &= (0,0,0,1,0,0,1,0,1,1)
\end{aligned}$$

The inner product between $f_1^{(3,2)}$ and f_H describes how much $f_1^{(3,2)}$ *lies in the direction* of H. Computing the inner products we get

	f_A	f_B	f_C	f_D	f_E
$f_1^{(3,2)}$	36.79	-2.21	-12.20	-8.21	-14.21

We observe that the first order effect lies most in the direction of A, being often in the minority with other voters, but lies least in the direction of E, being only occasionally in the minority of the pairs.

To interpret the second order effects, for each pair HK of X, define functions according to the criterion already explained, that is $f_{HK} \in M^{(3,2)}$ identifies the elements of $f^{(3,2)}$ which "contain" HK with 1 and the others with 0. So

$$\begin{aligned}
f_{AB} &= (1,0,0,0,0,0,0,0,0,0)\\
f_{AC} &= (0,1,0,0,0,0,0,0,0,0)\\
f_{AD} &= (0,0,1,0,0,0,0,0,0,0) \qquad \text{etc.}\\
f_{AE} &= (0,0,0,1,0,0,0,0,0,0)\\
f_{BC} &= (0,0,0,0,1,0,0,0,0,0)
\end{aligned}$$

Computing the inner products between $f_2^{(3,2)}$ and f_{HK} we get the exact data vector $f_2^{(3,2)}$

	f_{AB}	f_{AC}	f_{AD}	f_{AE}	f_{BC}	f_{BD}	f_{BE}	f_{CD}	f_{CE}	f_{DE}
$f_2^{(3,2)}$	0.17	2.50	4.17	-6.83	-0.50	-4.83	5.17	-1.50	-0.50	2.17

These results represents the **pure second order effects**, namely the pair's weight in the minority, after removing the mean effects and the effects of the individual. The values of

$$f_2^{(3,2)} = \begin{array}{cc} \begin{pmatrix} 0.17 \\ 2.50 \\ 4.17 \\ -6.83 \\ -0.50 \\ -4.83 \\ 5.17 \\ -1.50 \\ -0.50 \\ 2.17 \end{pmatrix} & \begin{array}{c} AB \\ AC \\ AD \\ AE \\ BC \\ BD \\ BE \\ CD \\ CE \\ DE \end{array} \end{array}$$

represent the pair's weight in the voting process.

We observe that the second order effect lies most in the direction of BE and least in the direction of AE.
Through this analysis we may point out *particular coalition effects* that do not arise from a direct analysis of data; for example, the pair DE has a quite high second order effect, whereas D and E have low values in the first order effects of the minority related to the individual ($f_1^{(4,1)}$). **This means that D and E are seldom in the minority alone, while they are often in the minority of the pairs.**

2.4 Analysis of the Supreme Court

In [22] Orrison and Lawson used noncommutative harmonic analysis to detect and analyze coalitions in the United State Supreme Court.

This application seems very interesting because shows the power of the machinery of generalized spectral analysis in detecting coalitions.

The United State Supreme Court has nine justices. So let $X = \{A_1, \ldots, A_9\}$ be the set of the nine voters. The function f is defined on k-element subsets of X, with $1 \leq k \leq 4$; this covers all cases since at most four justices can dissent on any given case. Unanimous cases are not considered.

Following the example of five people committee, if M is the vector space of all complex–valued functions on X, M decomposes into the direct sum of subspaces

$$M = M^{(8,1)} \oplus M^{(7,2)} \oplus M^{(6,3)} \oplus M^{(5,4)}$$

These subspaces decompose into invariant subspaces

$$\begin{aligned}
M^{(8,1)} &= M_0 \oplus M_1 \\
M^{(8,1)} &= M_0 \oplus M_1 \oplus M_2 \\
M^{(8,1)} &= M_0 \oplus M_1 \oplus M_2 \oplus M_3 \\
M^{(8,1)} &= M_0 \oplus M_1 \oplus M_2 \oplus M_3 \oplus M_4
\end{aligned}$$

The difference with the example of section 2.3 is that here we have in addition the projection on the subspaces M_3 and M_4; M_3 contains information about triples of justices and M_4 contains information about groups of four justices.

Orrison and Lawson analyzed voting of various Court: Warren Court from 1958 to 1962, Burger Court from 1967 to 1981, Renquist Court from 1994 to 1998. The analysis of these data is too long to be presented here; we refer to [22].

Nevertheless we want to report some considerations on the results of the analysis. Orrison and Lawson observed one similarity that goes through most of the Courts for the entire 1958-1998 period; it was the relationship of the second and fourth order effect in the decomposition of the 5-4 splits.
In most natural Courts the second order effect is the largest, followed by the four order effect. This suggest that when votes are close there are pairs of individual who tend to move together and also groups of four which tend to form. But single individuals and triples of individuals are less influential.
The large fourth order effect may represent the fact that the Supreme Court hears a number of cases on which they are closely divided (otherwise there would be no need to hear the case). It also may represent the fact that four justices must agree on the decision to hear a case.
The large second order effect is not so obvious, although it may represent the power that any two individuals can have in a closely divided case or possibly the ease with which to people cans strike a bargain to work together.

3
Spectral analysis and preference formation

In this Chapter we illustrate an application of generalized spectral analysis to preference formation.

Our context can be summarized as follows. We interpret the decision to vote for a particular party as a process of delegation to decision makers having a simplified system of preferences. Each person in a population votes for the political party that place priority on one or more issues that they consider important.

On the basis of a survey on preferences of population, we have simulated a delegation procedure which chart the selection process of a particular party. Making use of noncommutative harmonic analysis, we decomposed the delegation function and isolated the effect of a particular affinity, or a combination of either the pair of items that characterize a party. We used noncommutative harmonic analysis as an application of some results obtained by Michael E. Orrison and Brian L. Lawson in relation to spectral analysis applied in voting in political committees (see [22], [23] and [24]).

3.1 Introduction

Individuals facing a choice are often not able to make a full comparison between alternatives. Even if they are able to pin down their preferences for certain characteristics of an object (for instance, a car), they would probably be able to compare only a few of them. In the case of a car, one person would take into account room and safety, while somebody else's order ranking would be based on speed and acceleration. We can interpret this evaluation imagining that our "complete" selves delegate choices to a sort of simplified self.

Competition among products will be, in this way, not directed to the "real" population, but to the population of delegates that will choose products on the basis of a small subset of parameters. Car makers advertising speed and acceleration will not be considered by families who prioritize room and safety.

In public choice theory, political parties present themselves as decision makers committed to following a given preference order when faced with future choices. Parties collect delegations from people having similar preferences: in this way, instead of comparing all possible alternatives of the whole population, the number of alternatives is reduced to the number of parties. Traditionally this was intended in a similar way to the one used in economic location theory (see [13]). Parties have a complete system of preferences and they collect a delegation from the nearest people, i.e. from people having an order of preferences "not far" from the one expressed by the party.
Here instead of following this traditional path, we adopt a similar approach to the one presented in "car choice". **We describe parties as simplified systems of preferences and the process of delegation as giving the power of choice to parties that correspond to this simplified preference order**.

Given that parties compete to attract electors in a simplified preference space, the distribution of preferences will depend on the way preferences are simplified. If, for instance, parties simplify things proposing a couple of items to which they attach more importance, it could be that the items chosen complement themselves well, being able to attract a large share of voters, or alternatively the two items could reciprocally depress their power of attraction. When facing a simplified set of options, the right combination could be of fundamental importance.

In the following, first of all we present a general frame to formalize delegation over simplified preference orders. Then we illustrate a way to detect the "power of mixing" in a delegation procedure; this approach is based on the so–called noncommutative harmonic analysis or generalized spectral analysis (see [9] and [10]). Our work is an application of the results on spectral analysis of voting in committees due to Orrison and Lawson (see [22], [23] and [24]) explained in Chapter 2.

3.2 Individual preferences, parties and public choice

We begin by introducing some notations and terminology: a *party* will be defined as a simplified system of preferences, while the process of delegation to a party will be the power of choice corresponding to a simplified preference order.

Let X be a set of n objects. Let Λ be a set of m individuals. Λ will be called a **society** and the members of Λ **voters**.

Suppose that each individual of Λ is asked to **rank objects** of X putting them in a strict order, providing a total order on X.

Let Z be the set of all possible rankings over the elements of X; each $z \in Z$ may be viewed as a permutation of the n elements of X and each individual of Λ is asked to choose an element of Z.
We may define a total order on Z according to the choices of individuals of Λ, by counting the number of individuals that prefer each ranking. Let $z \in Z$, define β_z as the number of individuals of Λ choosing z. If $z_1, z_2 \in Z$, we define "$z_1 \leq_Z z_2$" if $\beta_{z_1} \leq \beta_{z_2}$ (where this last order "$\leq$" is the usual order on the natural numbers).
We call $(Z, \leq_Z)$ a set of **population preferences** over alternative rankings, according to the choices of individuals of Λ.

The set $(Z, \leq_Z)$ encapsulates the **individual preferences** arising from the society Λ; the theory of public choice (see [32] and [14]) allows us to define public choice functions which lead to a collective choice by starting from a collection of individual preferences.

DEFINITION 3.2.1 *Let $Z^\Lambda = \{f : \Lambda \longrightarrow Z\}$ be the set of functions from Λ to Z. Then a* ***public choice function*** *is a function $G : Z^\Lambda \longrightarrow X$.*

In other words, a public choice function associates a single ranking to each n-tuple of rankings which defines the consent of the population, according to some specified criterion.

In many cases it is difficult to obtain a public choice directly from the set of individual preferences, due to the large variety of possible preference orders. It is worthwhile then to look at "simplified" preference sets. In some sense, people **delegate** (see Vickers [44]) choices to delegates who have "similar" preferences. In political choices this is done by voting for a *party*.

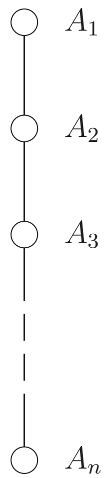
Let $X = \{A_1, \ldots, A_n\}$ be a set of n objects. Suppose that a total order $\leq_X$ is defined on X. We do not require that elements of X are strictly ordered through $\leq_X$, so obviously "=" may hold between some elements (this is different from the initial requests, according to which individuals of Λ are asked to order strictly the elements of X). Trivially $(X, \leq_X)$ is a lattice and may have a representation through lattice diagrams.

DEFINITION 3.2.2 *A* ***party*** *$\mathcal{P}$ may be defined by the preference order $\leq_X$, so $\mathcal{P}$ may be identified with lattice $(X, \leq_X)$.*

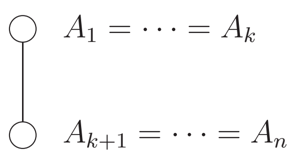
DEFINITION 3.2.3 *We define $\mathcal{P}$ a* ***complete party*** *if $\mathcal{P}$ is associated to a lattice $(X, \leq_X)$ where $\leq_X$ provides a strict order on the elements of X.*

DEFINITION 3.2.4 *We define $\mathcal{P}$ an* ***incomplete party*** *if $\mathcal{P}$ is associated to a lattice $(X, \leq_X)$ where $\leq_X$ provides a not strict order on the elements of X.*

An example of a complete party is

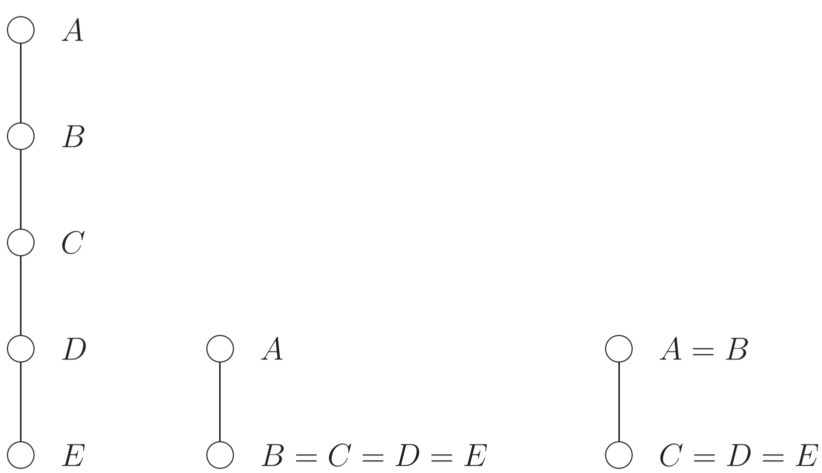


An example of an incomplete party is



with $k < n$.

For example, some preference orders over five objects A, B, C, D, E are



(a) (b) (c)

Parties like (a) are complete and parties like (b) or (c) are incomplete.

DEFINITION 3.2.5 *Let P be the set of all parties over X. We define a **party delegating** as a map*

$$\xi : Z \longrightarrow P \tag{3.1}$$

from the set of population preferences Z to the set of parties P.

We are going to investigate a party delegating map ξ empirically, in order to detect particular properties. We will decompose data related to ξ into many components, each of which will have specific meanings, according to the mathematical framework of generalized spectral analysis (see sections 1.5 and 2.2).

We observe that ξ in general does not satisfy order preservation; we recall that ξ preserves order if

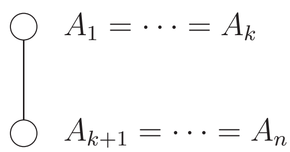
$$\text{for each } z_1, z_2 \in Z \text{ with } z_1 \leq_Z z_2, \text{ then } \xi(z_1) \leq_P \xi(z_2) \tag{3.2}$$

where the order on P is defined as on Z. In general it is not meaningless to have no order preservation: an individual may delegate a party even if it does not preserve his order of preferences; a distance may be defined between individual rankings arising from the choices of Λ and parties. ξ is then meaningful if it minimizes this distance, even if it does not preserve order.
As mentioned in the introduction, our approach is far from the traditional one used in economic location theory, where parties collect delegations from individuals having *similar* preferences. We describe parties as "simplified" systems of preferences and in the process of delegation an order preservation may be required. For this reason, we assume that ξ satisfies condition (3.2).

3.3 Voting for incomplete parties and the power of combination

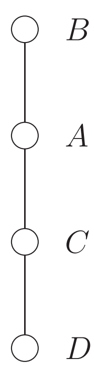
Let $\mathcal{P}$ be a complete party. An individual of society Λ votes for party $\mathcal{P}$ by the selection in Z of a complete ranking of the n objects of X.

Let $X = \{A_1, \ldots, A_n\}$ be a set of objects. Consider an incomplete party $\mathcal{P}_k$ of the form

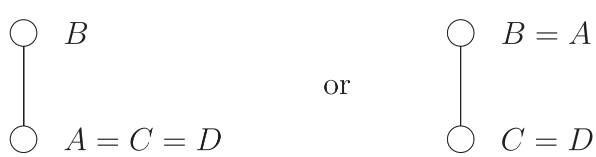


(3.3)

By selecting the incomplete party $\mathcal{P}_k$ the attention is focused on the first k alternatives chosen by an individual of Λ. A voter of Λ does not directly select an incomplete party, but proposes a complete ranking of preferences as described in section (3.2). If, for example, a person chooses the order $BACD$ over four objects, then this order corresponds directly to the complete party



But we may relate it also to the incomplete parties



and so on, according to which simplification we are dealing with; in this sense we talk about "simplification".

The approach through incomplete parties can be very useful if we suppose it is more straightforward and meaningful for a voter to concentrate on k alternatives, instead of n alternatives, with $k < n$.
In other words, by the delegation to an incomplete party as (3.3) the first k alternatives of a ranking are mixed together and act as a global single first choice, while the last $(n-k)$ items of the same ranking are also mixed and act as a global last choice. Obviously this interpretation of delegation to incomplete parties can be adapted to each type of incomplete party, not only to the example considered in (3.3).

3.4 Detecting the power of combination

In section (3.3) we focused our attention on simplified preference systems and their related incomplete parties. In this way a party earns consent in a reduced preference space and the distribution of preferences will depend on the way they are simplified. For example, suppose that an incomplete party is structured in such a manner that it proposes a pair of items as predominant: it could happen that these alternatives complement themselves strongly or alternatively they could weaken themselves reciprocally.

Orrison and Lawson (see Chapter 2 and [22], [23] and [24]) used generalized spectral analysis to locate influential coalitions in political voting procedures. We use generalized spectral analysis to detect what we call the *power of mixing or combination* in delegation procedures.

3.4.1 Preferences combination

Noncommutative harmonic analysis applied to an analysis of voting allows us to detect influential coalitions between voters of a committee (see Chapter 2).

The context we are dealing with proposes a set of voters Λ and a finite set $X = \{A_1, \ldots, A_n\}$ of alternatives to rank. In this setting it seems quite meaningless to look for influential coalitions between voters, because the society Λ can be composed by a huge number of members or by a sample of a population. We are interested in seeking a sort of "influential coalition" between preferences, even if this "dual" approach seems meaningless at this point.

Let $X = \{A_1, \dots, A_n\}$ be a set of n alternatives and suppose people of a society Λ is asked to rank $A_1, \dots A_n$, as prescribed in section (3.2). We refer to a notation of paragraph (2.2). Define

$$X^{(n-k,k)} = \text{ the set of } k\text{-elements subsets of the alternatives of } X$$

with $1 \leq k \leq \frac{n}{2}$.

Let $\omega \in X^{(n-k,k)}$, that is a set of k elements of X. Let $\mathcal{P}_{(k,\omega)}$ be the incomplete party corresponding to the lattice

"equality" on the elements of ω

"equality" on the elements of $X \setminus \omega$

For example, if $\omega = \{A_1, \dots, A_k\}$ is the set of the first k elements of X, then $\mathcal{P}_{(k,\omega)}$ has the form

$$A_1 = \cdots = A_k$$

$$A_{k+1} = \cdots = A_n$$

Define

$\gamma_{\mathcal{P}_{(k,\omega)}} :=$ the number of individuals of Λ who choose $A_1, \dots, A_k$ as their first k alternatives in their rankings (independently of the order) and $A_{k+1}, \dots, A_n$ as their last $n-k$ alternatives (independently of the order).

In other words, integer $\gamma_{\mathcal{P}_{(k,\omega)}}$ represents the number of individuals of Λ voting the party $\mathcal{P}_{(k,\omega)}$. Define a function $f^{(n-k,k)}$ on $X^{(n-k,k)}$ as

$$f^{(n-k,k)} := \gamma_{\mathcal{P}_{(k,\omega)}} \tag{3.4}$$

for each $\omega \in X^{(n-k,k)}$. We are interested in the spectral expansion of $f^{(n-k,k)}$, for each $1 \leq k \leq \frac{n}{2}$.

3.5 An application to a survey

We used the approach explained in section (3.4) to analyze some results of a survey on the preferences of the Trentino population. On the basis of the survey results, we simulated a delegation to hypothetical incomplete parties as defined in section (3.2). Moreover, making use of noncommutative harmonic analysis, we decomposed the resulting delegation function. In this way, the meaning of spectral expansion of the function defined in (3.4) will became clearer.

The **Indagine sulle preferenze della popolazione trentina**[1] (see [35]) is a survey carried out in 2004 on a sample of about 2000 adults resident in the province of Trento.

One of the research's aims was find out about the population preferences relative to some general themes of collective well–being; the knowledge of these preferences can be advantageously used to estimate potential impacts on the population of different types of public policies. One question in particular was useful for finding out about preferences:

Question n. 5 - collective well–being

In your opinion, what is more important between:

1. [A] full employment and [B] environment preservation?
2. [A] full employment and [C] health?
3. [A] full employment and [D] local income increase?
4. [A] full employment and [E] preservation of water and air quality?
5. [B] environment preservation and [C] health?
6. [B] environment preservation and [D] local income increase?
7. [B] environment preservation and [E] preservation of water and air quality?
8. [C] health and [D] local income increase?
9. [C] health and [E] preservation of water and air quality?
10. [D] local income increase and [E] preservation of water and air quality?

[1] We are grateful to Maurizio Pisati and Antonio Schizzerotto for giving us the accession to data of the survey *Indagine sulle preferenze della popolazione trentina*

Denote with

A	full employment
B	environment preservation
C	health
D	local income increase
E	preservation of water and air quality

Questions on collective well-being are structured as pairs comparisons between alternatives; such an approach may lead to preference systems that do not satisfy transitivity: for example, an interviewee may prefer A to B, B to C, but C to A. For our purpose it is meaningful to concentrate our investigation only on preference systems which satisfy transitivity. For this reason we established a simple way to detect if a preference system does satisfy transitivity or not. This method is based on simple considerations about matrices associated to preference systems.
We refer to preference systems satisfying transitivity as **consistent preference systems**.

We examined all the preference systems arising from question n. 5 of "Indagine" and established that they was inconsistent for a percentage of 34,2 %. For our analysis we used only data arising from consistent preference systems.

3.5.1 Transitivity of preferences

Let $X = \{A, B, C, D, E\}$ be the set of five alternatives investigated in question n. 5 of "Indagine". Each interviewee's answer can be realized as a table of the following type

Questions	1	2	3	4	5	6	7	8	9	10
Answers	B	C	A	A	C	D	B	C	C	D

where answer n. 1 stands for the choice between A and B, answer n. 2 for the choice between A and C and so on.

In general, the pairwise comparison adopted by the investigation of question n. 5 does not lead to a **consistent ordering** of all feasible alternatives. To make choices one needs only a choice function that allows one to select a best alternative from a set of possible alternatives. For example, an answer of type

Questions	1	2	3	4	5	6	7	8	9	10
Answers	B	C	A	A	C	D	B	C	C	D

does not lead to a total order of preferences, because a cycle between A, B and D exists, indeed B is preferred to A, A is preferred to D, but D is preferred to B. Conversely, an answer of type

Questions	1	2	3	4	5	6	7	8	9	10
Answers	A	A	A	A	C	D	E	C	D	D

satisfies transitivity and leads to the total preference order $CDEAB$.

Total orders

Let us recall some notations and terminology. Let $X = \{A_1, \dots, A_n\}$ be a set of n elements. Define on X a relation "$\leq_X$" (if there is no misunderstanding, we will use notation $\leq$) satisfying:

(A) reflexivity: $A_i \leq A_i$, $\forall i = 1, \dots, n$

(B) antisymmetry: if $A_i \leq A_j$ and $A_j \leq A_i$, then $A_i = A_j$, $\forall i, j = 1, \dots, n$

(C) comparability: for any $A_i, A_j \in \Omega$, either $A_i \leq A_j$ or $A_j \leq A_i$.

If "$\leq$" satisfies also

(D) transitivity: if $A_i \leq A_j$ and $A_j \leq A_k$, then $A_i \leq A_k$ $\forall i, j, k = 1, \dots, n$

"$\leq$" is called a **total order** on X.

We associate to $(X, \leq)$ a matrix which encapsulates the relation on X. Define $M_{(X,\leq)} := (m_{ij})$ where

$$m_{ij} = \begin{cases} 1 \text{ if } A_i \leq A_j \\ 0 \text{ otherwise} \end{cases} \qquad i, j = 1, \dots, n \tag{3.5}$$

We observe that $M_{(X,\leq)}$ satisfies the properties

$$i)\ \sum_{i,j=1}^{n} m_{ij} = \frac{n(n+1)}{2}$$

$$ii)\ m_{ij} = \begin{cases} 1 \text{ if } m_{ji} = 0 \\ 0 \text{ if } m_{ji} = 1 \end{cases} \qquad \forall i \neq j, \quad i,j = 1,\ldots,n$$

Property *ii)* is an obvious consequence of the definition of $M_{(X,\leq)}$. Property *i)* is a consequence of counting the number of 1 in the diagonal of $M_{(X,\leq)}$ plus the number of 1 appearing in the rest of the matrix, which is the number of possible unordered pairs of n elements.

PROPOSITION 3.5.1 *Let $X = \{A_1, \ldots, A_n\}$ be a set of n elements with a relation "$\leq$" satisfying (A), (B) and (C). Let $M_{(X,\leq)}$ be the matrix defined in (3.5). Then "$\leq$" satisfies transitivity (and in particular is a total order) if and only if, up to re-ordering the indexes of the elements of X, the matrix $M_{(X,\leq)}$ satisfies*

$$m_{ij} = \begin{cases} 1 \ \textit{if}\ i \leq j \\ 0 \ \textit{otherwise} \end{cases} \tag{3.6}$$

for all $i,j = 1,\ldots,n$, that is $M_{(X,\leq)}$ is strictly lower triangular of the form

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 1 & \cdots & 1 & 1 \end{pmatrix}.$$

PROOF. Suppose that "$\leq$" satisfies transitivity; in particular, it is a total order on X, so we may re-order the indexes of the elements of X so that $A_n \leq \cdots \leq A_2 \leq A_1$. Consequently:

$$1 = m_{11} = m_{21} = m_{31} = \cdots\cdots = m_{n1}$$
$$1 = m_{22} = m_{32} = m_{42} = \cdots = m_{n2}$$
$$1 = m_{nn}$$

and so $M_{(X,\leq)}$ has the desired form.
Conversely, suppose that $M_{(X,\leq)}$ satisfies (3.6). We want to prove that $A_n \leq \cdots \leq A_1$, so "$\leq$" is a total order on X and in particular satisfies

transitivity. We proceed by induction on n. For $n = 1$ there is nothing to prove. Suppose that $A_{j-1} \leq \cdots \leq A_1$. By hypothesis $m_{ij} = 0$ for each $i < j$. Then $A_j \leq A_1$, $A_j \leq A_2$, ..., $A_j \leq A_{j-1}$, for each $j = 1, \ldots, n$; by inductive hypothesis $A_{j-1} \leq \cdots \leq A_1$, then also $A_j \leq A_{j-1} \leq \cdots \leq A_1$. By induction we get $A_n \leq \cdots \leq A_1$. □

Proposition (3.5.1) does not provide a direct method to check if $M_{(X,\leq)}$ corresponds to a transitive relation. Nevertheless, an operative procedure can be easily found. Let $M_{(X,\leq)} = (m_{ij})$ be the matrix associated to $(X, \leq)$. Let $M^1, \ldots, , M^n$ be the column–vectors of $M_{(X,\leq)}$; obviously $M^i \in \mathcal{M}(n \times 1, \mathbb{R})$. Define

$$\alpha_j := \sum_{i=1}^{n} m_{ij} \qquad \forall j = 1, \ldots, n, \tag{3.7}$$

in other words α_j is the sum of the elements of column M^j in $M_{(X,\leq)}$. Observe that $\alpha_j \in \mathbb{N}$ and $0 < \alpha_j \leq n$, for each $j = 1, \ldots, n$. The following corollary is a different interpretation of proposition (3.5.1).

COROLLARY 3.5.1 *Let $X = \{A_1, \ldots, A_n\}$ be a set of n elements with a relation "$\leq$" satisfying (A), (B) and (C). Let $M_{(X,\leq)}$ be the matrix associated to $(X, \leq)$. Then "$\leq$" satisfies transitivity if and only if $\alpha_1, \ldots, \alpha_n$ can be "strictly" ordered.*

PROOF. Suppose that "$\leq$" satisfies transitivity. Then according to proposition (3.5.1)

$$M_{(X,\leq)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 1 & \cdots & 1 & 1 \end{pmatrix},$$

so $\alpha_1 = n, \alpha_2 = n - 1, \ldots, \alpha_n = 1$ and $\alpha_1 > \cdots > \alpha_n$.
Conversely, suppose that $\alpha_1, \ldots, \alpha_n$ can be strictly ordered. Suppose that, after re-ordering the indexes, $\alpha_1 > \alpha_2 > \cdots > \alpha_n$. Necessarily $\alpha_n = 1, \alpha_{n-1} = 2, \ldots, \alpha_2 = n - 1, \alpha_1 = n$, so

$$M_{(X,\leq)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 1 & \cdots & 1 & 1 \end{pmatrix}$$

and by proposition (3.5.1) "$\leq$" is transitive. □

EXAMPLE 3.5.1 *Let* $X = \{A_1, A_2, A_3, A_4\}$ *and*

$$M_{(X,\leq)} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We can see that $\alpha_1 = 3, \alpha_2 = 1, \alpha_3 = 2, \alpha_4 = 4$. We may associate to the α_js a vector $\alpha_{(X,\leq)} := (\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (3, 1, 2, 4)$ whose entries can be strictly ordered, so $\leq$ is a total order on Ω. In this case vector $\alpha_{(X,\leq)}$ provides also the order which is $A_2 \leq A_3 \leq A_1 \leq A_4$.

EXAMPLE 3.5.2 *Let* $X = \{A_1, A_2, A_3, A_4\}$ *and*

$$M_{(X,\leq)} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix}.$$

We see that $\alpha_1 = 3, \alpha_2 = 2, \alpha_3 = 3, \alpha_4 = 2$. The entries of $\alpha_{(X,\leq)} = (3, 2, 3, 2)$ cannot be strictly ordered, because some values are equal. According to corollary (3.5.1), $\leq$ does not satisfy transitivity. Actually, there is at least the cycle $A_1 \leq A_2, A_2 \leq A_4, A_4 \leq A_1$.

Consistent preference systems

The preferences expressed by the answers to question n. 5 define relations "$\leq$" on $X = \{A, B, C, D, E\}$ which satisfies reflexivity, antisymmetry, comparability, but not necessarily transitivity. Let $M_{(X,\leq_i)}$ be the 5×5 matrix associated to X and relation $\leq_i$ arising from the i-interviewee's answer. $M_{(X,\leq_i)}$ corresponds to a total order of preferences if and only if it satisfies the conditions of proposition (3.5.1) or equivalently of corollary (3.5.1).

Regarding question no. 5, the "Indagine" provided answers from the 1.898 people interviewed, but 54 of them gave incomplete answers; we are omitting such set of answers considering only **1.844** interviews. We associated a matrix $M_{(X,\leq_i)}$ to each i-interviewee's answer and analyzed the results of

“Indagine” using techniques of corollary (3.5.1), seeking **consistent preference systems**, that is, preference systems which satisfy transitivity. We found out that **1.213** preference systems are **consistent**, against **631** which are not, for a percentage of **34,2 %** of inconsistent systems. Among the consistent systems, we found the distribution of orders illustrated in table (3.1), in which the number of people who choose some order is associated to each chosen order.

ABEDC	1	CADBE	18	CEABD	189	EADCB	1
ACBDE	1	CADEB	35	CEADB	90	EBACD	1
ACBED	5	CAEBD	97	CEBAD	210	EBCAD	14
ACDBE	1	CAEDB	48	CEBDA	49	EBCDA	4
ACDEB	9	CBADE	4	CEDAB	28	ECABD	43
ACEBD	7	CBAED	20	CEDBA	24	ECADB	19
ACEDB	7	CBDEA	4	DABEC	1	ECBAD	49
ADCBE	1	CBEAD	64	DCABE	1	ECBDA	15
ADCEB	1	CBEDA	17	DCAEB	1	ECDAB	6
BCAED	1	CDABE	8	DCEAB	1	ECDBA	6
BCDAE	1	CDAEB	17	DEACB	1	EDABC	1
BCEAD	6	CDBAE	2	DECBA	1	EDBAC	1
BECAD	7	CDBEA	2	EABCD	2	EDBCA	1
BECDA	2	CDEAB	7	EACBD	2	EDCAB	1
CABDE	18	CDEBA	6	EACDB	1	EDCBA	3
CABED	30						

Table 3.1: Consistent preferences systems

We observe that not all the possible orders on the five elements A, B, C, D, E have been chosen; the possible consistent preferences systems on 5 alternatives are $120 = 5!$ whereas only 61 have been selected.

3.5.2 Generalized spectral analysis

Let $X = \{A, B, C, D, E\}$. According to notation of paragraph (2.2) we have

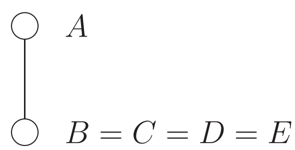
$$X^{(4,1)} = \{A, B, C, D, E\}$$
$$X^{(3,2)} = \{AB, AC, AD, AE, BC, BD, BE, CD, CE, DE\}$$

where notation AB stands for the subset $\{A, B\}$. Let $f^{(n-k,k)}$ be the function defined in (3.4). In our context

$$f^{(4,1)}(A) = \gamma_{\mathcal{P}_{(1,A)}}$$

that is

$$f^{(4,1)}(A) = \text{the number of people choosing the incomplete party}$$



We need to count the incomplete parties with 1 alternative in the *first position*, arising from the data of "Indagine". In table (3.1) of Appendix B we show how the consistent preference systems from "Indagine" have been chosen by the interviewed people. We rewrite table (3.1) as

A	33
B	17
C	987
D	6
E	170

Table 3.2: Parties with 1 predominant preference

where the number of people choosing an order with A *at the first place* or B or C and so on is pointed out . So we have

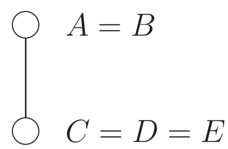
$$f^{(4,1)} = \begin{pmatrix} 33 \\ 17 \\ 987 \\ 6 \\ 170 \end{pmatrix} \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array}$$

In the same way

$$f^{(3,2)}(AB) = \gamma_{\mathcal{P}_{(2,AB)}}$$

that is

$$f^{(3,2)}(AB) = \text{the number of individuals choosing the incomplete party}$$



For this data, table (3.1) becomes

AB	1	BD	0
AC	276	BE	28
AD	3	CD	45
AE	6	CE	728
BC	117	DE	9

Table 3.3: Parties with 2 predominant preferences

where the number of people choosing an order with A and B *at the first two positions* **independently of the order**, or choosing A and C and so on, is pointed out. We have

$$f^{(3,2)} = \begin{array}{cc} \begin{pmatrix} 1 \\ 276 \\ 3 \\ 6 \\ 117 \\ 0 \\ 28 \\ 45 \\ 728 \\ 9 \end{pmatrix} & \begin{array}{c} AB \\ AC \\ AD \\ AE \\ BC \\ BD \\ BE \\ CD \\ CE \\ DE \end{array} \end{array}$$

First order effects

As outlined in the example of section 2.3, we may project the function $f^{(4,1)}$ onto the invariant subspaces of the decomposition $M^{(4,1)} = M_0^{(4,1)} \oplus M_1^{(4,1)}$ and get the Fourier expansion of $f^{(4,1)}$:

$$f^{(4,1)} = \begin{pmatrix} 33 \\ 17 \\ 987 \\ 6 \\ 170 \end{pmatrix} = \begin{pmatrix} 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \end{pmatrix} + \begin{array}{c} \begin{pmatrix} -88.30 \\ -104.30 \\ 865.70 \\ -115.30 \\ 48.70 \end{pmatrix} \end{array} \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array}$$

where

$$f_0^{(4,1)} = \begin{pmatrix} 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \end{pmatrix}, \qquad f_0^{(4,1)} = \begin{pmatrix} -88.30 \\ -104.30 \\ 865.70 \\ -115.30 \\ 48.70 \end{pmatrix}$$

The function $f_1^{(4,1)}$ shows the **first order effect**, that is the amount which each party of the form $\mathcal{P}_{(1,A)}$ differs from the mean. In this case the interpretation of the first order effects does not yield new information in relation to the initial data $f^{(4,1)}$.

Parties proposing 1 alternatives	first order effects
C = party: health	865.70
E = party: preservation of water and air quality	48.70
A = party: full employment	-88.30
B = party: environment preservation	-104.30
D = party: local income increase	-115.30

Table 3.4: First order effects of parties with 1 alternatives

Second order effects

We may project the function $f^{(3,2)}$ onto the invariant subspaces of the decomposition $M^{(3,2)} = M_0^{(3,2)} \oplus M_1^{(3,2)} \oplus M_2^{(3,2)}$ and get

$$f^{(3,2)} = \begin{pmatrix} 1 \\ 276 \\ 3 \\ 6 \\ 117 \\ 0 \\ 28 \\ 45 \\ 728 \\ 9 \end{pmatrix} = \begin{pmatrix} 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \end{pmatrix} + \begin{pmatrix} -179.47 \\ 160.53 \\ -209.13 \\ 28.87 \\ 113.87 \\ -255.80 \\ -17.80 \\ 84.20 \\ 322.20 \\ -47.47 \end{pmatrix} + \begin{pmatrix} 59.17 \\ -5.83 \\ 90.83 \\ -144.17 \\ -118.17 \\ 134.50 \\ -75.50 \\ -160.50 \\ 284.50 \\ -64.83 \end{pmatrix} \begin{array}{c} AB \\ AC \\ AD \\ AE \\ BC \\ BD \\ BE \\ CD \\ CE \\ DE \end{array}$$

where

$$f_0^{(3,2)} = \begin{pmatrix} 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \\ 121.30 \end{pmatrix}, \quad f_1^{(3,2)} = \begin{pmatrix} -179.47 \\ 160.53 \\ -209.13 \\ 28.87 \\ 113.87 \\ -255.80 \\ -17.80 \\ 84.20 \\ 322.20 \\ -47.47 \end{pmatrix}, \quad f_2^{(3,2)} = \begin{pmatrix} 59.17 \\ -5.83 \\ 90.83 \\ -144.17 \\ -118.17 \\ 134.50 \\ -75.50 \\ -160.50 \\ 284.50 \\ -64.83 \end{pmatrix}.$$

The function $f_2^{(3,2)}$ represents the **second order effects**, that is the weight of a party which proposes two alternatives, after removing the mean effects and the first order effects.

Parties proposing 2 alternatives			second order effects
CE = party: health	+	water-air	284.50
BD = party: environment	+	income	134.50
AD = party: employment	+	income	90.83
AB = party: employment	+	environment	59.17
AC = party: employment	+	health	-5.83
DE = party: income	+	water-air	-64.83
BE = party: environment	+	water-air	-75.50
BC = party: environment	+	health	-118.17
AE = party: employment	+	water-air	-114.17
CD = party: health	+	income	-160.50

Table 3.5: Second order effects of parties with 2 alternatives

Interpretation: Mallow's method

We use Mallow's method (see section 2.3.3) to interpret the first and second order effects $f_1^{(3,2)}$ and $f_2^{(3,2)}$.

To interpret the first order effects we compute the inner product between $f_1^{(3,2)}$ and the "naturally interpretable" functions defined in section 2.3.3, that is

$$\begin{aligned} f_A &= (1,1,1,1,0,0,0,0,0,0)\\ f_B &= (1,0,0,0,1,1,1,0,0,0)\\ f_C &= (0,1,0,0,1,0,0,1,1,0)\\ f_D &= (0,0,1,0,1,0,0,1,0,1)\\ f_E &= (0,0,0,1,0,0,1,0,1,1) \end{aligned}$$

We get

	f_A	f_B	f_C	f_D	f_E
$f_1^{(3,2)}$	-199.20	-399.20	680.80	-428.20	285.80

Table 3.6: First order effects

Table (3.6) shows that the first order effect *lies mostly in the direction* of C and least in the direction of D, which confirms the results for parties proposing one alternatives of table (3.4).

To interpret the second order effects we compute the inner product between $f_2^{(3,2)}$ and the related "naturally interpretable" functions

$$\begin{aligned} f_{AB} &= (1,0,0,0,0,0,0,0,0,0) \\ f_{AC} &= (0,1,0,0,0,0,0,0,0,0) \\ f_{AD} &= (0,0,1,0,0,0,0,0,0,0) \qquad \text{etc.} \\ f_{AE} &= (0,0,0,1,0,0,0,0,0,0) \\ f_{BC} &= (0,0,0,0,1,0,0,0,0,0) \end{aligned}$$

and get

	f_{AB}	f_{AC}	f_{AD}	f_{AE}	f_{BC}
$f_2^{(3,2)}$	59.17	-5.83	90.83	-144.17	-118.17

	f_{BD}	f_{BE}	f_{CD}	f_{CE}	f_{DE}
	134.50	-75.50	-160.50	284.50	-64.83

Table 3.7: Second order effects

which is exactly the vector $f_2^{(3,2)}$. We observe that the second order effect *lies in the direction* of pairs

CE = party: health + water/air preservation
BD = party: environment + income
AD = party: employment + income

This means that there is an *intrinsic affinity* between pairs for which the second order effect is high. We try to explain this concept. According to the first order effects of $f^{(4,1)}$ of table (3.4), two items have a particular value, C and E respectively, which correspond to the incomplete parties

(a_1) (a_2)

According to the second order effects of $f^{(3,2)}$ of table (3.7), the powerful pairs are CE, BD and AD respectively, that is the incomplete parties

$$\begin{array}{c} C=E \\ | \\ A=B=D \\ \\ (a_3) \end{array} , \qquad \begin{array}{c} B=D \\ | \\ A=C=E \\ \\ (a_4) \end{array} \quad \text{and} \quad \begin{array}{c} A=D \\ | \\ B=C=E \\ \\ (a_5) \end{array}$$

are winning. This means that C and E are powerful items either alone (that is when a party proposes one predominant alternative) or together (that is when a party proposes two predominant alternatives). This is not the case of parties (a_4) and (a_5); for example, B and D are weak alone, but they get stronger in a pair. The same for A and D.
The analysis of second order effects allows to understand if there are "intrinsic affinities" between items and if coupling items contributes to weaken or strengthen them.

Orrison and Lawson, in their spectral analysis of voting data of the United State Supreme Court (see [23]), suggest to display the results of the second order effects analysis in a graphical way which helps to single out particular "coalition effects".
This approach can be applied in our context. Figure (3.1) displays information of table (3.7) in a way that makes it easy to identify the "coalition effects" arising from the second order effects analysis.

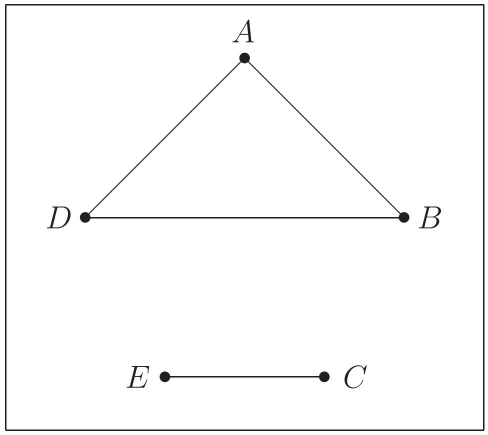


Figure 3.1: Second order effects

Two items are joined by a line when a positive second order effect exists between them; the absence of a line means that the second order effect of the pairs is negative.

Significance

One of the main idea of harmonic analysis (both commutative, that is the classical spectral analysis, or noncommutative) is to find subspaces into which data can be decomposed, while preserving the most important structure of the data, as explained in section (2.2), but also to see *which subspaces contain the largest amount of the data*. This is done by considering the "length" of the vectors arising from the decomposition, in order to determine which vector is significant.

The traditional method for determining **significance** is to compare the norm squared of the f_i vectors divided by the dimension of the subspaces M_i. As suggested by Diaconis (see Diaconis [10] pag. 954 and also Orrison and Lawson [22]), for the type of data we are examining this may be misleading; it is better to consider the norm squared of the vectors. In our situation we have

$$||f_1^{(3,2)}||^2 = 294420.94$$

$$||f_2^{(3,2)}||^2 = 173074.77$$

Comparing these two values suggests that **the first order effects are more significant than the second order effects**.

3.6 Conclusions

In this Chapter we have interpreted the way people vote for parties as a process of delegation to decision makers using a simplified system of preferences. Moreover, on the basis of a survey on preferences of the population, we have simulated a delegation procedure to parties. Finally, making use of noncommutative harmonic analysis, we decomposed the delegation function, and isolated the effect of affinity, or mixing, between the pair of items that characterize a party.

This approach appears to be promising both to understand how people, with limited rationality, act given a simplified set of options, and to empirically study the best way to simplify a preference set, in order to gain from complementarity among objects over which people express an order of preference.

Further studies should be devoted to enlarge the model. A first approach could be an extension of the voting model to a general choice among objects, in the line of the introductory example. As it was stated by [40], "if one replace the term 'individual i' with 'property i', social choice theory is transformed from a theory of social decision into a theory of formation of individualistic preferences" (p. 58). In this direction, we argue that our frame could be used to model a process of choice under limited rationality assumptions, where agents are unable to evaluate all the characteristics of goods, defined as in [20], and to compare them with their complete preferences.

A second approach could be directed to refine the supply side of the model. In our model we assumed parties as given; but in fact there could be a competitive formation of parties. They could in fact choose to be more or less specialized, proposing a shorter or longer list of characterizing items. An empirical analysis of the kind we described in this paper could help to ex ante define the best positioning of parties.

Bibliography

[1] U. Baum, *Existence and efficient construction of Fast Fourier Transform on supersolvable groups*, Comput. Complexity, **1**, (1991), 235–256.

[2] U. Baum and M. Clausen, *Fast Fourier Transform for symmetric groups: theory and implementation*, Mathematics of Computation, **61** no. 204, (1993), 833–847.

[3] L. Beckett and P. Diaconis, *Spectral analysis for discrete longitudinal data*, Advances in Mathematics, **103**, (1994), 107–128.

[4] M. Beckmann, *Location theory*, Random House Inc., New York, 1968.

[5] M. Clausen, *Fast Generalized Fourier Transforms*, Theoretical Computer Science, **67**, (1989), 55–63.

[6] C. Chatfield, *The analysis of time series: an introduction*, Chapman & Hall, New York, 1987.

[7] J. W. Cooley and J. W. Tukey, *An algorithm for machine calculation of complex Fourier series*, Mathematics of Computation, **19**, (1965), 297–301.

[8] D. Coppersmith and S. Winograd, *Matrix multiplication via arithmetic progressions*, J. Symbolic Comput., **9** no. 3, (1990), 251–280.

[9] P. Diaconis, *Group representations in probability and statistics*, IMS, Hayward, CA, 1988.

[10] P. Diaconis, *A generalization of spectral analysis with application to ranked data*, The 1987 Wald Memorial Lectures, The Annals of Statistics, **17** no. 3, (1989), 949–979.

[11] P. Diaconis and D. N. Rockmore, *Efficient computation of the Fourier Transform on finite groups*, J. Amer. Math. Soc., **3** no. 2, (1990), 297–332.

[12] P. Diaconis and D. N. Rockmore, *Efficient computation of isotypic projections on the symmetric group*, eds. L. Finkelstein and W. M. Kantor, DIMACS Series in Discrete Math. and Theoret. Comput. Sci., AMS, Providence, RI, **11**, (1993), 879–104.

[13] J. J. Gabszewicz, J. F. Thisse, M. Fujita and U. Schweizer, *Location theory*, Harwood Academic Publishers, Chur, 1986.

[14] P. C. Fishburn, *The theory of social choice*, Princeton University Press, 1973.

[15] E. M. Hoover, *The location of economic activity*, McGraw Hill, New York, 1948.

[16] W. Isard, *Location and space-economy*, MIT Press, Cambridge, Massachusetts, 1956.

[17] A. Izeman and S. Zabell, *Babies and the blackout: the genesis of a misconception*, Technical Report no. 38, Dept. of Statistics, University of Chicago, 1978.

[18] G. D. James, *The representation theory of the symmetric groups*, Springer–Verlag, Berlin, 1978.

[19] G. D. James and A. Kerber, *The representation theory of the symmetric group*, Addison–Wesley, Reading, Massachusetts, 1981.

[20] K. J. Lancaster, *A new approach to consumer theory*, Journal of Political Economy, **74** no. 2, (1966), 132–157.

[21] K. J. Lancaster, *Variety, equity and efficiency*, Columbia University Press, 1979.

[22] B. L. Lawson and M. E. Orrison, *Analyzing voting from a new perspective: Applying spectral analysis to the U.S. Supreme Court*, Presented at the annual meeting of the American Political Science Association, Boston, August 29, 2002.

[23] B. L. Lawson, M. E. Orrison and D. T. Uminsky, *Discrete Analysis of Voting in Committees*, Presented at the annual meeting of the Midwest Political Science Association, Chicago, April 3–6, 2003.

[24] B. L. Lawson, M. E. Orrison and D. T. Uminsky, *Noncommutative Harmonic Analysis of Voting in Small Committees*, July, 2003.

[25] C. L. Mallows, *Non–null rankings models I*, Biometrika, **44**, (1957), 114–130.

[26] D. K. Maslen and D. N. Rockmore, *Separation of Variables and the Computation of Fourier Transforms on Finite Groups. Part I*, J. Amer. Math. Soc., **10** no. 1, (1997), 169–214.

[27] D. K. Maslen, *The efficient computation of Fourier Transforms on the symmetric group*, Mathematics of Computation, **67** no. 223, (1998), 1121–1147.

[28] D. K. Maslen and D. N. Rockmore, *The Cooley–Tukey FFT and group theory*, Notices of the AMS, November, 2001.

[29] D. K. Maslen, M. E. Orrison and D. N. Rockmore, *Computing the isotypic projections with the Lanczos iteration*, SIAM J. Matrix Analysis and Application, **25** no. 3, (2004), 784–803.

[30] D. K. Maslen and D. N. Rockmore, *Adapted diameters and the efficient computation of Fourier Transform on finite groups* in *Proceedings of the sixth annual ACM-SIAM symposium on Discrete algorithms*, San Francisco, California, (1995), 253–262.

[31] D. K. Maslen and D. N. Rockmore, *Generalized FFTs. A survey of some recent results* in *Groups and computation II*, eds. L. Finkelstein and W. M. Kantor, DIMACS Series in Discrete Math. and Theoret. Comput. Sci., AMS, Providence, RI, **28**, (1997), 183-237.

[32] D. C. Mueller, *Public choice II.* A revised edition of *Public choice*. Cambridge University Press, Cambridge, 1989.

[33] N. A. Naimark and A. I. Stern, *Theory of group representations*, Springer–Verlag, New York, 1982.

[34] M. E. Orrison, *An eigenspace approach to decomposition representations of finite group*, Ph.D. Thesis, Dartmouth College, 2001.

[35] M. Pisati, *Benessere collettivo e preferenze della popolazione: un'analisi empirica*, Quaderni della programmazione PAT, *Metodi e applicazioni di ricerca valutativa per la pubblica amministrazione*, Trento, 2004.

[36] J. N. Rayner, *An introduction to spectral analysis*, Monographs in spatial and environmental system analysis, Pion Limited, London, 1971.

[37] S. Rensi and E. Zaninotto, *Simplified preferences, voting, and the power of combination*, preprint, Trento, 2004.

[38] D. N. Rockmore, *Fast Fourier Transforms for Wreath Products*, Appl. Comput. Harmon. Anal., **2** no. 3, (1995), 279–292.

[39] D. N. Rockmore, *Some applications of generalized FFTs*, in *Groups and computation II*, eds. L. Finkelstein and W. M. Kantor, DIMACS Series in Discrete Math. and Theoret. Comput. Sci., AMS, Providence, RI, **28**, (1997), 329–369.

[40] A. Rubinstein, *Economics and Language. Five essays*, Cambridge University Press, Cambridge, 2000.

[41] B. E. Sagan, *The symmetric group. Representations, combinatorial algorithms, and symmetric functions*, Springer–Verlag, New York, 2001.

[42] J. P. Serre, *Linear representations of finite groups*, Springer–Verlag, New York, 1977.

[43] G. L. Thompson, *Generalized permutation polytopes and exploratory graphical methods for ranked data*, Ann. Stat., **21** no. 9, (1993), 1401–1430.

[44] J. Vickers, *Delegation and the theory of the firm*, Economic Journal, **95**, (1985), 138–147.

[45] A. Weber, *Theory of the location of industries*, Russell & Russell, New York, 1971.